

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 8, 2004, 12:03:42 ; Search time 34846 Seconds
(without alignments)
11772.891 Million cell updates/sec

Title: PCT-US04-13989-18

Perfect score: 8675

Sequence: 1 aaccggcgtaactggggggtg.....ctggcaccgcgcatgtctct 8675

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 4526729 seqs, 23644849745 residues

Total number of hits satisfying chosen parameters: 9053458

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl:*
1: gb_ba:*
2: gb_hgt:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_scs:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2797.6	32.2	6815	1	AB076662 Pantoea a
2	2688.4	31.0	6586	1	AY166713 Pantoea s
3	2618	30.2	6918	1	ERWCR7 Erwinia ure
4	2616.4	30.2	6918	6	I12910 Sequence 13
5	2486.2	28.7	6965	1	ERWCR7A Pantoea agg
6	2248	25.9	12753	1	ERWCR7S Pantoea agg
7	1885.6	21.7	4624	6	CQ793208 Sequence
8	1345.8	15.5	349107	1	BX571873 Phototrab
9	1265.8	14.6	2415	1	ERWCR7B Pantoea agg
10	1156.4	13.3	3801	1	AF408848 Xanthobac
11	1085.8	12.5	6635	1	AF218415 PMCR7WYI
12	1054.6	12.2	6335	1	ATUCRTWA
13	1053.6	12.1	5373	1	ATUCRTWA
14	1032.2	11.9	5188	1	PAU62808
15	1032.2	11.9	8625	6	AR169831 Sequence
16	1032.2	11.9	8625	6	AR452005 Sequence
17	1032.2	11.9	11233	6	AR169852 Sequence
18	1032.2	11.9	11233	6	AR452026 Sequence
19	826.2	9.5	1479	6	CQ793426 Sequence

20	826.2	9.5	1479	6	CQ793595	CQ793595 Sequence
21	826.2	9.5	1479	6	E22381	E22381 beta-Carote
22	826.2	9.5	1479	6	I12907	I12907 Sequence 10
23	826.2	9.5	1493	6	BD241820	BD241820 Method of
24	826.2	9.5	1493	6	AX014697	AX014697 Sequence
25	811.8	9.4	1479	6	AX394988	AX394988 Sequence
26	751.4	8.7	1518	6	I72653	I72653 Sequence 7
27	746.4	8.6	1522	6	I72654	I72654 Sequence 9
28	655.2	7.6	349723	1	BX842650	BX842650 Bellolovib
29	611.4	7.0	1482	6	AS8565	AS8565 Sequence 6
30	611.4	7.0	1482	6	AB84702	AB84702 Sequence 12
31	611.4	7.0	1482	6	AR103081	AR103081 Sequence
32	611.4	7.0	1482	6	AR139697	AR139697 Sequence
33	611.4	7.0	1482	6	AR391876	AR391876 Sequence
34	562	6.5	1198	6	I24838	I24838 Sequence 5
35	562	6.5	1198	6	I72652	I72652 Sequence 5
36	504.4	5.8	1083	6	I40224	I40224 Sequence 1
37	483	5.6	349737	1	BX572597	BX572597 Rhodosphe
38	466.2	5.4	30040	1	AF182374	AF182374 Bradyrhiz
39	456.6	5.3	1296	6	AX394984	AX394984 Sequence
40	452.2	5.2	1235	6	I72656	I72656 Sequence 13
41	449.8	5.2	5397	12	AT605097	AT605097 Synthetic
42	441	5.1	1235	6	I72657	I72657 Sequence 15
43	440	5.1	1149	6	AX394986	AX394986 Sequence
44	429	4.9	1581	1	D83514	D83514 Erythroba
45	429	4.9	1581	6	E11103	E11103 Gene coding

ALIGNMENTS

RESULT 1	AB076662	6815 bp	DNA	linear	BCT 26-DEC-2001
LOCUS	AB076662	Pantoea agglomerans pv. milletiae ctxE, ctxH, ctxY, ctxI, ctxB, ctxZ genes, complete cds.			
DEFINITION	AB076662				
ACCESSION	AB076662.1	GI:18143444			
VERSION	AB076662.1	GI:18143444			
KEYWORDS					
SOURCE					
ORGANISM					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
REFERENCE					

```
gene
1081..2376
/ gene="crlx"
1081..2376
/ gene="crlx"
/ codon_start=1
/ protein_id="BAB79601.1"
/ db_xref="GI:18143446"
/ translation="MSHFAVIAPPEYSHVOALQHLISQALIRGHQITFIQOTDVSALL
TDSRIGFEPGLASHPAGSLAHTLQLAHPLGSMKLIEMARSTDMCRELPVIVS
TLDIDVINDOMERBAGLAALNLPVSVACALPIAREDPPLAVMPPEXPATDQAR
ERTTSKITYDMLRRHRDVIYARVAMMGLAPREKHCPSVAQINQLPEIDFPO
TLPDHEFVSGLPRTTEBPAPASOPRYEPHGDPRPIFASLGLQHRGLKTTIVACQ
EIDAOULAHGCRISPQAEKLAQASHVQVDPADQAAALAOQLDVLTHGGMVTDG
KHLRLPLTPIPLAFDQGVARVVMGIGRRASFTTSHMAQLQTLADESVQDM
KTRSLRQAGGTIADIVIEQMLTRQPVLTRHYAAV"
2363..3523
/ gene="crlx"
2363..3523
/ gene="crlx"
/ codon_start=1
/ protein_id="BAB79602.1"
/ db_xref="GI:18143447"
/ translation="MPRYDLILVAGLNGULALRLRQSPRLILLIDAREBRGANH
TWSFABDLTETQHRMIAPIVHVHMPGEYERFORSRSMNGYFCVTAERFAOYIRR
FAPDLINTRVVASIASRAVTLDDGRLVESAVIDRGYOPDGLRMGFOSFOCEOML
SEBGLFAPILIMATVDOQAGYRFVSLPESADTLIEDTHYIDNLTEDGRARONLH
DYAOCQGMRLDRLRERQALPITLGDVAAPFQONRLPESGLRAGLPHPTTYSYPL
AYVALDRLAQMOTFTSETLHATITQPRASQMOORFRPMRNMLFLAGPADQKQVQO
RFFGLPREGLIARFAGKTLPLRLRLISGKRPVVALAQLAITPHRQAMO"
3520..4998
/ gene="crlx"
3520..4998
/ gene="crlx"
/ codon_start=1
/ protein_id="BAB79603.1"
/ db_xref="GI:18143448"
/ translation="MNRITVIGAGFGSLALRILOASGVPTRLLEORDKRGAAVYVO
DQGPFPAGPTVITDPSAIRELPLTAKGKISDYLVEPVKPYFRALCMESGKVSND
QPALBAOITAPNPRDVBGRRPFLAISAIVABGTLKGTVPISFPMRLAAQOLAKL
QARSVYSKVAASYIEDHELRQAFSPHSLVGNPFAVSYLYTLIHAIERWGVFPFG
GTGALVOGWKLPEDLGEVLELNASVRLTQENRITVAHLKQGRVETRAVANSNDY
VHTYRELISQHPASQAQGRSLQNMKMSNLFVIYFGILHNNOLAHTTVEGFRYSEL
IDELFNKDGIAEDFSLYLAFCVTDPSLAEGGSGYVLAIPVHLGTAIDMAVEGPR
LRRIPIPEOHMPPGRLSOLVTHRIPTPDPDELNAVQGSAPSVEPILITQAMFRP
HNDDKINNTYLVGAGTHPGAGIPGVIGSAKATAGMLBDA"
5034..5924
/ gene="crlx"
5034..5924
/ gene="crlx"
/ codon_start=1
/ protein_id="BAB79604.1"
/ db_xref="GI:18143449"
/ translation="MEVGSKFATASKLFGAKTRRSYMLYAMCRHCDVDVLDQVLGF
SNDTPISQSAEQRLOLEMKTRQAVASQMBEPAFAFCVAAHNDLIPYADHLAG
FAMDVHETRYOTLDDTLRCYHAGVAGVMAOIMGRDVAATLDRACDGLAOLNPI
ARDIVEDADAGRCYLPATWLAEBGLTRENADONKALSRVARLVEETPEPYRBSAS
AGLPGULPRLSAMAIATIQVYRKIGMKVQVQASQAMDOROSTSPEKALLVAASQA
VTSRVNARHARSADLQRPV"
complement (5863..6390)
/ gene="crlx"
complement (5863..6390)
/ gene="crlx"
/ codon_start=1
/ protein_id="BAB79605.1"
/ db_xref="GI:18143450"
/ translation="MLMIMNALLIVLVIVIGMEITALLAHRYIMHGCMGWHLSHNEBH
```

ORIGIN									
Query Match		32.2%;	Score 2797.6;	DB 1;	Length 6815;				
Best Local Similarity		66.5%;	Pred. No. 0;						
Matches 4140;		Conservative	0;	Mismatches 2054;	Indels 30;	Gaps 8;			
QY	1243	AAAGCATCTCTTCAGTTGCAATTCAGCAATTCGAGGCCCATCTTGAACATTTATTCGCTG	1302						
DB	200	ACAGCCGATGACGCCAGCCTCTTGAACGACATTTAGCAACCGCTTGATCACTTTTACCGG	259						
QY	1303	CCGGGACAGACGCCGATCGGCTTCGGGCGCCCATCTGCTGCCGACGCTTGGCGACACCGGGCA	1362						
DB	260	TTGAAAGCGAACGTGACTTATGTTGGGCGCTGCCATTCGCCAGCGGTGGCTGGCACAGGA	319						
QY	1363	AACGTAATTCGCTCGCTCTTGTCTGCTGCGACAGCCGATATGGGCTGTGACGTGGCGC	1422						
DB	320	AGCGTATCCGTCCACTGCTGTTGTTGGTCAGCGCGCATCTGGGCTGCAACGCCACGC	379						
QY	1423	AGCAGGGCATCTTGAATCTTGTGCTGTGCGGTCGAAATGTCGACGCTGCTCATCTGATCC	1482						
DB	380	CTGCCGCGCTGCTGATCTGCTGCGCGGTGAAGATGGTGATGCGCATCTGATTC	439						
QY	1483	TGCAACGACATTCATGATCAATGATTAAGCCCGGATGCGAGCTGGAGCGCCCGGCAATCCACT	1542						
DB	440	TGGAATGACATGCCCTGATGATGATGATGACGAACTGCTGCCGAGACGTCCGACCATTCATT	499						
QY	1543	GTGAATPAGGGGAAAGCGTGCGCATCTGTGACAGCGGTGCGGCTACTCAGCGCGCTTTG	1602						
DB	500	GCCAGTATGATGAAACATGTGCGCATTTCTGGCGCGGCTGCGGCTGCTGATGAAGCATTCG	559						
QY	1603	AGGTGATTCGCTTCGCGCGCGGTCTGCGACAAACGACAAAGCCGATTCGCGAGC	1662						
DB	560	GCGTGTGCTGCGGCGCAAGAGCTTAAACGCAACCCGACAGCGCATCTGTCGACGAAT	619						
QY	1663	TCTCTCTGCGCGTGGGCGCTCAGAGGACTGTGTTAGGGTCATGTTCCAGGATCTGATGACG	1722						
DB	620	TATCCACGAGATCGGATGATGACAGGGCGTGTGACGGGCAATTTAAGATCTCTCGAGG	679						
QY	1723	GCGCACACAGCGCGATGCGGAAAGCCATCACTGACCAATGAACTGAAACCAAGCGTCC	1782						
DB	680	GTGACAAAGCCACGACGAGCGCTGACGCGCATTTGATGACCAATCACTAATAACAGCACGC	739						
QY	1783	TGTTTCGCGCAGCGTGCAGATGCGCGCGCATTTGCGGCGCATGCGTACAGTACGTC	1842						
DB	740	TGTTTCGCGCTCCATGACGATGAGCTCTATCTGCTGTAAGCTCAGGTGAAGCGCCCG	799						
QY	1843	AGCGTTAAGCTATTTTGGCGCAGAGATTTAGTCAGAGCTTCCAGTTAATGACAGACCTGG	1902						
DB	800	AACAGCTGACACGTTTTTGTGCTTAATCTTGTTAGGCTTTCAGAGCTATGAGAGATCTCA	859						
QY	1903	CGGATGCGCTTAAACACACACCGGACAGAGACTGTGATCAGAGATCAGGGCAATTCACGCTGG	1962						
DB	860	CTGACCGCATGAGCGCGACACCGGTAAGATGCCATCAGATGACGGGAATCAACGCTGG	919						
QY	1963	TGCAAGTCTGGGCGCGGAAAGGGGCTGAGCGTCTGTGCGCGCATCTTAAGACGGCGG	2022						
DB	920	TGAATCTGCTGGAGACACAGCGGCTTGAAACGCGACCTGCGCGCATCTTCGCGTGGCCA	979						
QY	2023	ATGCACACCTTGGCTGCGCTGCGCATCGCGGTCCGCCACCCGTCMAATATGACAGGCC	2082						
DB	980	GCGAGCATCTGTTATGCGGCTGCGCAGGACGGTTATGCAACACCAATTTGTTCAAGCCT	1039						
QY	2083	TGTTTAAATCAACAGCTGGCGATTTCACTGAAGCGGCTCATCTATGGGCGATTTTGGC	2142						
DB	1040	GGTTTGAAGAAAACTGCTGCGCGCTCATGTTAAGAGTCT-----GATAGGCCACTTTTGC	1094						
QY	2143	CGTTATTCGCGCACCGCTTACAGCACTTTCACGATTCGACGCGCGCTGGCGAAACGCT	2202						
DB	1095	GGTCATTGACCGCGCTTCAACAGCATGTCAGGGCGCTTACAGCACTTTCAGCGCGCTT	1154						

QY	2203	GCTGGCGCGGGAATGCGATCGATCACTTTATTCAGCAAAATGATGACGCACTTGTCTGAG	2262
Db	1155	AATGGCAACGCGGAATCAAGATTACTTTTCATCAGACAGCGAAGCTTAGCGGCTACTCAC	1214
QY	2263	CGACGAGCGCATGCGCTTTGTGGCCGTGGGAGAGGACAGCATCTCGGCCGATCGCTCTC	2322
Db	1215	CGATAGCCGTAACGGCTTTTTCGCCGTGGTTTAGCTCGGATCCGGGGGCACTTGGC	1274
QY	2323	CAGCGAATCTCAGGGGCGCTGGCCGGAACCGGGGGGGCTGTCCGTGTTTCGCGATTCACGA	2382
Db	1275	ACACACCCCTGGAATCTGGGGGGGCAATCCCTCTGGCCCCCTCAATGCTGAACACTGATCAATGA	1334
QY	2383	TCTGGCCAGCAACCAACGATATGCTATGCGCGAACTGCGCGGTGCTGCACGACTGCA	2442
Db	1335	GATGGCCCGCAGCACCGGATATGCTCTGTGCTGAACCTGCCCGTGGTGTACACACCTTGGC	1394
QY	2443	GGTGATGGCGGTGATGGCCGGAATCCAAATCGAAGCGGCTGTGTGTCTGTGTGGAGAGGCTT	2502
Db	1395	GATTGATGGCGGTGATTTGTGATACAGATGGAACCTGGCCGGTCACTGGGGGAGAAAGGCT	1454
QY	2503	ACAGCTGCGCTTCGATGTCGTGGGCGCTGGCGGCTGGCCGATCAATGCGCAAGGGGCATTTCC	2562
Db	1455	GAACCTGCTTATGTTTGGTTGCTGTGCTGTGGCCTGCACTTAAACGTGAAGCGAATTTCC	1514
QY	2563	GCTGGTGTGATGCCCTTTTCGCTTTGCTCAGGAATGAGAAAGCGGTGACGCTATTCAGGC	2622
Db	1515	GCTGGCGGTGATGCCGTTTGAATTTGACAGGACCGATCAGGCAAGGAAAGCTATTCGAC	1574
QY	2623	CAGCAGTGAATCTTACGACCGGCAATCATGCGTGTCTATGAGCGCTCATTCGTCTCATGC	2682
Db	1575	CAGTGAATAAATCTATGATGCTGGCTGATGGACGTCACATCGGGTGTATCGGCCGCAACGC	1634
QY	2683	GCGGCGCTTCGGGCGTGGCCGGAAGCGAATGGCTTACATCAGTGTGTGTCGCGCGCTGCGCA	2742
Db	1635	TCAATGCCATGGGAATTAGACACACGGGAAAACTGATCATCTGTTTTCACGGGTGGGCA	1694
QY	2743	AATCAGTCACTGTGGTCCCGCTTTTGAATTTTTCACGCCAGCAACTGCCAGCCTGCTATCA	2802
Db	1695	GATCAACACAGCTGATTCGGGAATGGAATTTTCCCGCAGACACTGCCCGCATTTTCCA	1754
QY	2803	CAGGTGGGTCCGCTGGGAC-----TCCAGTTGTAGCGGCGGCGTGGCCGACCTTGG	2857
Db	1755	TTTCGGTTGGCCCGGTGGGTACACTGAACCGGACCCGCGCATCTCAGCTCCCTACTT	1814
QY	2858	CCAGCGGCTGGCGCAG-CCGGTGTGTATGCTCGTGGGAGACGGTACAAGGGGCAATCCGTT	2916
Db	1815	CCCGCATGGCACAAACCGCGCAATTTTGTCTCACTCGGCAACGCTGCAGGAGCAATCGTA	1874
QY	2917	TGCGCTGTTTTCGATCTGCTCAGGCGCTGCAGCAATCAGCAGCTGTGCTGTGTGTGGC	2976
Db	1875	CGGCTGTGTTTAAACATCGTCCGGCGCTGCAGGAAGATGATGCTCATGTTATTTGGTGGC	1934
QY	2977	ACACTGTGGCGGGTTTGAACCGCCAGCCAGGCAATCAGTCAAGCTGGCCGGTGTGGCTG	3036
Db	1935	GCAATTTGGGTGCTGTGTCACTTTCCAGGCGGAATACTGGCGAGGCGACGCACTGTTC	1994
QY	3037	GGTACCGGATTTTGTGATTCAGCGGGGGGGCGCTGAGATGCGGAACGTTTTATCATCTCA	3096
Db	1995	GGTGTGTGATTTTTCGACCAAGGGGGGCTGCACCTGGCGAGGCGCACTTTGTCAATTAAGCA	2054
QY	3097	CGCGGCTCTGAACAGTGGCGCTGGAAGCACTGAGTGTGGCAACCGCGGAATGCTGGCGCTGC	3156
Db	2055	CGGGGGGATGAATACGGTGTGGATGGGATTTAACCACTGACGGCCGTGCTGAAGATCTCC	2114
QY	3157	GATGCGCTTGCATCAGCCCGGCGTGGCGGCAAGTATGATGGCAACGGCGTGGCCGAGCG	3216
Db	2115	GCTGGCTTTTGTATCAGCCGGGTGTGGTGGCCAGAGTGTCTGTGGCAACGGTATTCGAGCTGC	2174
QY	3217	CGGCTCAACGTTTCAAGCGGGGTGGGAGCTGGAGACAACCTGGAACAGTTGTGTAAGTA	3276
Db	2175	CGCCTCAAGCTTTCACACAGCACTCTCATGTGACGTAGCTTCAAGACTTATCTGGCTGA	2234
QY	3277	CGATTCGCTATGCTGCGCATGTGACGCATTCAGGCGGACGCTGCAAGGGGCGGCTGGCTG	3336

Db	2235	TGAAGCTACGTTACGACGATGAAACCGCTGCGTAGCGCACTTGTCTACGGGGCGGAAC	2294
QY	3337	TACGGCGCGGCTGATATATGTCGAGCAGGGCGCTGTGTCAGACAGAAATCGTGTGCGGA	3396
Db	2295	CACCTGGCGCGGACATGTCGAAACAGGCGATCTGAACCCGACCGCGGTCTCACCAAG	2355
QY	3397	GGCCACCTGATGCGCGCACCTTATGATGTCAATTCGTGTCGTGCGCGGCTGGCTAAACGG	3456
Db	2355	GAGACACTATGTCGGG---GTATGATCTGATTTCTGGTAGGTGCGGGACTGGCCAAACGGG	2410
QY	3457	CTGATTTGGCTGTGCTTTACGCCAGCTGACGCCGCACTTAAAGTTTGTCTACTGAGAGT	3516
Db	2411	CTGATTCGCGCTCGCCCTGAGACACACAGCGGCCCTCCCTGCGCACTTCTGCTATTTGACCC	2470
QY	3517	CAGGGGACCGCGCGCGCAATCAACCGTGTTCATGCGGAAAGCGTAGGAAACG	3576
Db	2471	GAACTGTAACCGGGTGCACATCAACCTGGTCGTTTCACTGCGAAAGATCTACCGAAGCG	2530
QY	3577	CAGTTTCGCTGTGCTGAGCGCGCTCTTTCGCGCGCTGACCCTGGTTATCAGATACGCTTC	3636
Db	2531	CAGCATGCGCTGGATGCTGCTCGCTGTGTATATATCACTGCGCTGTATAGAGTCGCTTT	2590
QY	3637	CCCAACCTGCGTCCGACGCTGATGTGTGAATATTTGCTGATTCCTCGAGAAATTTGCC	3696
Db	2591	CCCCAACGCAATCCGACGATGAACAGTGGCTATTTTTCGTAATGCGGACGCGCTTCGCG	2650
QY	3697	CGGCACCTTAACGACAGGTGTCGTGTCGCGCGCTAGACCGACGCGCGCTGACGGAAGTCT	3756
Db	2651	CAGGTCACTCCGCAAGAGTTTGGCCCGGATTCGTCTGTAATACCGGGTTCGACGATCT	2710
QY	3757	TCACCAACCGGGGTGACACTGTGCGGATGTGCGGATGTTCACAGCGCGCGCGTATTTGAC	3816
Db	2711	GCTTCAGCGCGGTCACGCTGTGAAGATGGCCGGGTGCTGAGAGTGAACGCAGTATTTGAT	2770
QY	3817	GGACGCGGGCTGCAGCCGACACCGCATCTGAGCTCGCTATCAGGCAATTTGTGTGTAC	3876
Db	2771	GCGCGTGGCTACAGCCCGATGGCGCGCTGCGCATGGGCTTTCAGTCTGTTGTGTGACAG	2830
QY	3877	GAGTGGCAATGGCGCGCGCGGACGCTGAGACGCAATTTGATGAGACGCGACGCTC	3936
Db	2831	GAGTGGCACTCAGCGAGCCACATGGCTCGACCGCGCGATCATATGAGATGCCACGCTA	2890
QY	3937	GATCAGCAGCAGGGTTATTCGCTTTGTTTACACCTGCGCTCAGTGTGCGACGCTTTACTG	3996
Db	2891	GATCAGCAGGAGGCTATCGCTTTGTCTACAGCTTGCCTTTTCTGACGACACCTGTGTC	2956
QY	3997	ATTGAAGATACCACTACATCAACCATGTCCACGCTGATGTCGACAGGCGCGCTTAC	4056
Db	2951	ATTGAAGATACCACTACATGATTAACGCCACGCTGGAAGGCGATGCGCGCTGTACAGAT	3010
QY	4057	ATTACGATTTATGCCCACACAGCGGCGCTGGAATTTGGCCAGCGTGTGCGGAGAGACAC	4116
Db	3011	ATTCACGACTATATGCGCGCAGACAGGATGGGATTTAAGCCGCTGTACGCGGAAGGAA	3070
QY	4117	GGCTGCGCTCGCGATCAGCGTCAGGGGCGAATTCGATCAGTTCTGTGGCAACAGACAGCGG	4176
Db	3071	GGCGGTTTACCATTTAGCTGAACCGGAGCGTGGCGGCTTTCTGG---CAGAACCGGAT	3127
QY	4177	CAACCGTGCAGCGGAGCTGCGCGCGCGAATGTTTTCAGCGCACACCGGTTACTGTGTCGC	4236
Db	3128	TTGGCCGTGACGCGGTTTGGCGTGGCGGGTTATTCATCCCACTACCGGCTATTTCTGTGCA	3183
QY	4237	GGCGCGGTGGCGCTGGCGGAGAAATTTGCCAGACGCTGTCCCGCGGACGCTCAACGCTG	4296
Db	3188	CTGCGCGGTGCTCTGGCTGACCGGCTGGCGCGATGACAGATTTTAACTCTGGAACCTGT	3247
QY	4297	AGCCACTGCATCAATCTTTGGCCGTCAGCACTGGCGCGAGCGCTTTTTCGCTGTG	4356
Db	3248	CACGCCACCAATTACGCAATTTGCGCAATGACGCTGTGGACGAAACGCGCTTCTTCCGATG	3307
QY	4357	TTAATGCGATGCTGTCTTTCTTTCGGAACGCGCTGAACAGCGCTGGCGGTATGCAAGCT	4416

Db	3308	CTGAACCGTAAAGCTGTTTCTGTGGCGGGTCCGGCTACACAGCGCTGGCAGGTTATGCAAGCT	3367
QY	4417	TTTTACCGGCTTGAAGCGCCGAGTTGATTAAGCCGCTTTTAACGCCGGGCAACTGCGCCTCAAGC	4476
Db	3368	TTTTTAACGGCCTTCCCGAAGGTCTATAGCCCGCGTTTATATGGGGAAAACTCACTTTCCT	3427
QY	4477	GATTAAGACGCACTTCGTGTGGGGAAACCGCGGGTCCCTTCGGCGAAAGGCTCGGCGCA	4536
Db	3428	GATGCGCTGCGCACTTTAAGCGGGCAGCCCCGGTCCCCGTTCTGGGGGGCTTACAGGC-	3486
QY	4537	TTGATGATGACCTCTCCGTTACAGGGGAAATATGAAAAGCACTTATGTATGTGCGC	4596
Db	3487	--TATTATGACTCTCCACCTGTCAACAGCGCGATGCAATGAAATAGAACTCAAGTATTTGGCG	3544
QY	4597	CAGGCTTCGGTGGCGCTGGCGGCTGGCGAATGTCGTGCAAGGGGCGGGGTGGCGGTCAAGC	4656
Db	3545	CAGGCTTGTGGTGGTGGCTCTGGCAATTCGCTTCAGGCGTCAAGGGCTTCCACCCGAC	3604
QY	4657	TGCTGGAACACAGCGGATTAAGCCTTGCGGGCGCGCCTATGTGTATCAGATCAGGTTTAA	4716
Db	3605	TGCTGGAGCAACGTGAACAAGCTGGCGGGCTTATGTCTATCAGGATCAGGACTTCA	3664
QY	4717	CCTTTGATGCGCGGTCCAGCGGTATTAACCATCCCAAGGCTTATGAGGCGCTGTTTAAAGC	4776
Db	3665	CGTTTGAAGCGGGGCCCAAGGTATCACCGATCCAGGCGCATTTGAAGAGCTGTTCACCC	3724
QY	4777	TGGCAGGCAACAACTCAGTATTAATGTGCACTGATCCGGTACAGCCATTTTATCGC	4836
Db	3725	TCCGCGGTAAAAAGTCTCTGATCTATGTGCAAGCTGATGCCGTGTAAGCCGTTTATCGCC	3784
QY	4837	TGTGCTGGGAAGACGGCAGGCACTGTGACTACAGACAATCAGCGCAGCTGAGCAAGC	4896
Db	3785	TCTGCTGGGAATCCGGCAAGGTGTCACTTATGACAAAGATCAGCCCGCTGGAAAGCGC	3844
QY	4897	AGATTGGCACTTTTAAATCCCAAGATGTGCGCGGTTACCGCGAGTTTGTGCGCTATTAC	4956
Db	3845	AGATTGCCCGCTTATCCGCTGACAGCTTGAAGATATCTGTGCTTCTGGCGCTATTCCC	3904
QY	4957	AGGATGTGTTTCGTAGGGGCTATCTGAAACTGGGCAAGCTTTCGTGATTTCCGGC	5016
Db	3905	GAGCGGTGTTTGCTGAAGGCTATCGAAGCTTGGCACGTCGCTCGTTTCTGATATCCGCG	3964
QY	5017	ACATGCTGCGTCCGGGCGCACAGCTGGGGTCCGGCTGCAAGCCCTGGGGCAGTGTCAACGA	5076
Db	3965	ACATGCTCGGGCGCGCTCAAGCTGGCAAAACTTCAGCGGTGGGCGACGGTTTACAGCA	4024
QY	5077	TGATGCGCAAAATTTATCATGACATCACTGCGCAGGCTTTTCTTTCACTCGTTGC	5136
Db	4025	AAGTGGCAGCTACATTGAAGTATGAGCATCTGCGTCAAGGCTTCTCTTCCACTCATCTGC	4084
QY	5137	TGCTGCGCGGTATCTCTTTGCAACGCTTGGATTTAACTTAACTTAACTTAACTGAGC	5196
Db	4085	TGTGTGGCGGAATCCGTTTGGCACTTCTCAATTAACCTGATTAACCTGATCGCTGGAAC	4144
QY	5197	GCGAATGGGGCGTGTGTTTCGCGCGCGCGGATACCGGTGCGCTGTTGATGAGATGCGC	5256
Db	4145	GTTGAATGGGGCGTCTGTGTTCCCGCGCGGTGGGACCGGGCGCTGTGTGAGGGAGTGTGA	4204
QY	5257	GCGTGTTCGCGATTTGGCGGGTGAAGTCTGTCTCAACGCCGAGATCAAGCACTGTGAGA	5316
Db	4205	AACCTTTTGAAGATCTGGGCGCGCGAAGTGAAGTCAATGACAGCCTTCCCGGCTGGAGA	4264
QY	5317	CCGAGGTTAACCGGATCAAGCGGTGTCCAGCTGAAGATGGGGGCGCGTTTGGCGCGCGCG	5376
Db	4265	CCAGAGAAAAAGAATTACCGGGGGTCACTGAAAAGTGGCCGGGATCTTCCGACCCGCG	4324
QY	5377	CGTTTGCCTCAAAATCTGACGTGTGTGATCACTTACGATCGCCTGTTAAGCAAGATCTGTG	5436
Db	4325	CGTTTGCCTCAAAAGATGTGTGTTCAACTTACCGGAATGCTGTGAGCCAGATCCCG	4384
QY	5437	CGGCGCGTAAATCGCGGGCAACGCTGAAGCGTAAGCGTAAGCAACTCGCTGTTTGAAC	5496
Db	4385	CTTCCAGCGCGGAGGACGATTACTGTCAAAACAAACGATGAGCAACTCACTGTTTGTGA	4444

QY	5497	TCATATTTGGCTTAATCAATCAACCCGAGCGGGGACCAACAGGTGGCTTGGTGC	5556
Db	4445	TCATATTTGGCTTAATCAATCAATCACTGGGGCACACAGCTGGCTTGGTGC	4504
QY	5557	CGCGCTATCGGAATTGATCGATAGATCTTTCATATGACGACAGTGGCGGAATTTCT	5616
Db	4505	CGCGCTATCGGAATTGATGATGATGATCTTTCATTAAGAATGGCGCTGGCAGAGGACTTCT	4564
QY	5617	CGCTGTATCTGCATATGGCCCTTCCTCCAGCGATCCGTGGCTGGGACCGCGGGGGCTGGGCA	5676
Db	4565	CGCTGTATCTGCATATGGCCCTTCCTCCAGCGATCCGTGGCTGGGACCGCGGGGGCTGGGCA	4624
QY	5677	GTTTTCATGTCCTGGCGCGGTGGCGGATCTCGGTACCGCGCAATTCGACATGCGCAAGG	5736
Db	4625	GCTACTAGTGTCTGGCGCGCAATTCGCGACCTCGGACCGCGGATATCGACTGGGCGGTGG	4684
QY	5737	AAGGGCGCGCTTGGCGGATTCGCAATTTTGTCTTCTGGAAGAGACTATATATGGCGGGTC	5796
Db	4685	AAGGTCGCGCGCTCGCGCATCGCAATTTTGTCTTCTGGAACAGCACTATATATGGCGGGTC	4744
QY	5797	TGCGACAGCACTTATGTGACACACCGTATGTTTACGCGGTTGATTTTCGCGACACGCTGC	5856
Db	4745	TGCGTACGCACTTGTCTACGATTCGATCTTTCACGCGGTTGATTTTCGCGATAGCTGTA	4804
QY	5857	AACGCGATCAGAGGCTCAGCGTTTTGCTCGAACCCATTTTGAAGCAAAACGCGCTGGTTCC	5916
Db	4805	ATGCGTATCAGAGGTTCCGCGCTTCTCGGGAGAGCGGATCTTCGCAAAAGCGCTGGTTCC	4864
QY	5917	GGCGCGATTAACCGGATGCCGACATTTACTACTTTATCTGGTGGGGGCTGGGACGCGATC	5976
Db	4865	GGCTTCACAAACCGGATTAATAATATATATATCTCTATCTGGTGGGGGCTGGGACGCGATC	4924
QY	5977	CCGAGTCCGAGTGTGACAGGCGCTGATCGGCTTCGCGAAAGCGACCGCCAGCTGATGGTGC	6036
Db	4925	CTGGCGGGGGATTTCCAGCGGTAATTTGGCTCGGCAAGGCTTACCGCAGGATTAATGCTGG	4984
QY	6037	AGGATCTGACCGGATGAACCAACCGCGCTGATTTGAGCAGGTTCACGCAACCATATGGCGCA	6096
Db	4985	AGGATCTG---GCTTGAAATAGTCCGTCACTGCTGATCATGCGGTAGACACCATATGAGGT	5041
QY	6097	GGGCTCCAAAAGTTTCGCGACGCGTACCGGCTATTTGATCTTTCACACGCGCGCAGTAC	6156
Db	5042	GGGATCGAAAAGCTTTGCGCACCGGCTCAAAACTGTTTGGTCCAAAACCGACCGAAGGT	5101
QY	6157	GCTGATGCTGTAACGCTGTGTGTCTGCTACTGTGACGATGTATATGCTCAGACGCTGGG	6216
Db	5102	GCTGATGCTTAACGCTGTGTGTCTGCTACTGTGATGATGTGATTAACGATTCAGTATCTGGG	5161
QY	6217	CGAAGGCGGACGCGACGACGCGGGTGGCGGATTCGACAGGCGGGGATGGCGCCCTGCGAAAT	6276
Db	5162	ATTTCGACAAAGATACGCGCATGCTGCGAATCCGCGACAGCGGCTGGCGACGCTGGAGAT	5221
QY	6277	CGAAACCCGCGCGGCTTACAGCGGTGGCCACATGGATGAACACAGGTTTCGTGCTTTCA	6336
Db	5222	GAAAACGGGTCAAGGCTATGCGCGATTCGCAATGGATGATGAGCGCGGCTTTGACGCTTTCA	5281
QY	6337	GGAATGGCGCTGACGCACTCAAGCTTCCGACAGCTGGCTTTTGTATCATCTGGAAGGTT	6396
Db	5282	GGAGGTGGCAATGGCACAGATATTTCCCTGCTTACCTTTTATATCTGGGGGGGCTT	5341
QY	6397	TGGCATGGAATGGCGGTGAAGAACTTATGCGGTTCGAGGACACAGCTGCTTACTGCTA	6456
Db	5342	TGGCATGGAATGGCATGAGACAGCTTATCAAGCGTGGATGATATACGCTGGCTTACTGTTA	5401
QY	6457	TCACGTGGCCGCGCTGATGGGGTTAATGATGGCGCGCTGATGGGCGTACTGATGACGG	6516
Db	5402	CCACGTGGCGGCGGTGTTGGCGCTGATGATGGCGGAGATTTATGGGCGTATCCGACACGCG	5461
QY	6517	CGTACTCGATCAGCGCTGTGATTTGGCTCTGCGGTTTACGCTTTCGAAATATTCGACGGGA	6576
Db	5462	CACGCTGATGCGCGCTCGGATCTCGTCTGGCGCTTTCAGCTGACCAATATTCGCGCGCA	5522


```
/protein_id="AA085599.1"
/db_xref="GI:27228293"
/translation="MKPTIVAGFGGLALAIRLQAGIPVLLIEQRDKRGRAVYVQ
EQGFIDAGPVIIDPSAIEELFALAKQKDYVELLPVPEYKLQWSESKVFNVD
QAQLEAQIQOFNPRVDVAGYRAFLDYSAVNEGLKGTVPFLSKMLAAAPOLATL
QAMRSYKVAQYIEDEHLROAFSFLVAGNEPKLGSITLTLHALEREMGWFPFG
GTGLVNGMKLFDGLGEEVLYNARVSHMEVGDKIOAYOLEDRERETCAVSNMAY
VHTYRDLISQHPAAKQAKYLSKRMNSLPVLYEGNHHDDLAHHTVCFGRPRRL
IHEIFNMDLAEDSLYLHAPCVDPDSIAPGCGSYVLAPVPHLGTANLDMVAGEPR
LNRDIFDLYQHYMPGLRSQLYTHRMPTPDPLDELNAWQSGASVBPILITQSAWFR
HNRDKHIDNLYLVAGTHPGAGIPGVIGSAAKATGLMLBDLI"
5043..5933
/gene="crtB"
/locus_tag="5043..5933"
/gene="crtB"
/codon_start=1
/translation="phycoene synthase"
/protein_id="AA085600.1"
/db_xref="GI:27228294"
/translation="MAVGSKSFATSLFPDKTRRSYLMLYAMRCHDDYIDDTLGF
HADPSSQMPBQRLQLEMKTRQAYSQMHEPAFAFQEVAMAHDTAPAYADHLEG
FAMDVRETRLTLDLRLCYCHAVAGVLMMAQIMGRDADNATDRACDLGLAQITNI
ARDIVDAQVRCYLPESMLEBEGITANTAYAAAPENRQALSRIGRLVREAPYVSSM
AGLAOLPLRSAMAIATAKOYRKIGVKEQAGKQAMDRQSTAEKLTLLTASGA
VTSRMKTYPPRPAILMORPI"
complement(5872..6399)
/gene="crtZ"
complement(5872..6399)
/gene="crtZ"
/codon_start=1
/translation="beta-carotene hydroxylase"
/protein_id="AA085601.1"
/db_xref="GI:27228295"
/translation="MLMINALIVFVTVGMEVVAALAHKYLHMGWGMHLSHHEPR
KGAENVDLAAVFAIVSIALYFGSGTPIWGIQMGAYGLVYMWVDAVHVOHW
PFRYIPRKYLKRLYMAHRMHNAVRGECVSRFLYAPLSKLQALRLRHRMARSA
ARBDQDQVDVTSRSGK"

ORIGIN
Query Match 31.0%; Score 2688.4; DB 1; Length 6586;
Best Local Similarity 65.3%; Pred. No. 0;
Matches 4077; Conservative 0; Mismatches 2136; Indels 27; Gaps 8;

QY 1257 CAGTTCATACGCGATTCAGAGCCCATCTTGAACATTATTGCTGCCGGCAGACAGCC 1316
DB 220 CAGTTGCTGCTGATATCGATAGCCGCTTGATCAGTTACTGCGGTTCAAGGGTGAACGG 279
QY 1317 GATCGCGTTCGGGCGCGCATGCGCGGCGGACGCTGGACCGGGCAACGATTCGTCGG 1376
DB 280 GATTGTGTGGGTGCGCGATGCGTGAAGGACGCTGGACCGGGCAACGATTCGTCGG 339
QY 1377 CTCTTGCTGCTGCGGACGACGCGATATGGGCTGTGACGTGGCGACGAGGCGATCCTT 1436
DB 340 ATGCTGCTGTTATTAACAGCGCGCGATCTTGCTGTGCGATCAGTCACGGGGAGTTACTG 399
QY 1437 GATCTTCCTGTGGGCGTGAATGTGTGACGCTGCTCACTGATCTTCGACGACATTCGA 1496
DB 400 GATTTACCTTCGCGGCTTGAATGTGTGATGCTGCTGCTGCTGATTCGATGATATGCC 459
QY 1497 TCATAGATTAACGCGCGGATGCGACGCTGGGGCGCGGCAATTCACCTGTGAATATGGGAA 1556
DB 460 TGCATGACGATGCGGACGATGCGTGGGGGCGTTCACACATTCACACGCGAGTACGGTAA 519
QY 1557 AACGTGCGATCTCTGACGCGGTGCGCTACTACGCGCGCTTTGAGTGATTCCTC 1616
DB 520 CATGTGCGATTCGTGGCGGGGTGCTTTACTACGAAAGGTTTGGGGTGAATGGCGAG 579
QY 1617 GCGCGGGGTCTGCCAGACGACAAAGCGAAGCCATTGCCAGCTTCTCTGCGCG 1676
DB 580 GCTGAAGTCTGACGCGGATGCGAAACTGCGCGGCTGTGCGAGCTGTCACTGCGATT 639
```

```
QY 1677 GGCTGAGGGAAGCTGTGTCAGGGTCACTTCCAGAGATCTGATGACGCGGACACAGCCG 1736
DB 640 GGAATGAGGGGTCTGTTTCAGGGCGAGTTTAAAGACCTTCGGAAGCGATTAACCCGCG 699
QY 1737 AGTCGGAAACCATCAACCTCGAACCAATGAAGTAAACAGAGCTTCCTGTTCCGCCACG 1796
DB 700 AGCGCGATCCCATACGTCGTCACCATCAATCAATGATTTAAACAGACAGCGCTTTTCCGCTCA 759
QY 1797 CTGCAATGCGCGGATTCGCGCGCATGCGCTCAAGTCAAGTCAAGCTTAAGCTAT 1856
DB 760 ACCCAATAGCGTCCATTCGCGGCAACGCGCTCGGGAAGCGCGTGAACCTGCATCGT 819
QY 1857 TTTCGCGAGGATTTAGGTCAGGCTTTCGAGTTTCTGAGACGCTGGCGGATGCGCTCTAG 1916
DB 820 TTCTCGCTCATCTCGCGAGGCGCTTTCAGTTGAGTTGACGATCTTACCGATGCGATGACC 879
QY 1917 CACACCGGCAAGGACTGTCTACAGATTCAGAGGCAAAATCCACGCTGTGTGACATGTGGG 1976
DB 880 GATACCGGCAAGACATCAATCAAGATGACAGTAAATCAACGCTGGTCAATTTATTAAGC 939
QY 1977 CCGAAGGGGCTAGCGGTGCTGCGGACCATCTTAAGACGCGCCGATGACACCTTTGCC 2036
DB 940 TCAGCGCGCGGTGGAAGACCGCTGCGACAGCATTTGCGCTGCGCAGTGAACACTTTCC 999
QY 2037 TGGCGCTGCATGCGG---TGTGCGCAACCCGCAATATATGACGCCCTGTTAATCA 2093
DB 1000 GCGGCAATGCCAAACGGCCATTCCACCAACCACTTTTATTCAGGCTGTGGTTTGACAA 1059
QY 2094 CAGCTGCGCATGTTCAACTGAAGCCGCTCATACCTATGAGGSCATTTTGGCGTTATTCGC 2153
DB 1060 AAACGTGCGCGCTCAGTTAAGATGCT-----GANTAGCCATTTTGGCGTATGCGAC 1114
QY 2154 CAGCGCTCTACAGCCACTTTCACGCAATTCGAGGCGCTGGCGCAACGCTGCGCGCG 2213
DB 1115 CGCGCTTTTACGCCATGTTCCGCTGCTGCAAAACCTTGCTCAGAAATTAAGTGGCCGCG 1174
QY 2214 GACATGCGATCACTTTATACAGAAAGATGATCAGACGACCTTGCTGAGGACGAGGCGA 2273
DB 1175 GTATGCTGTTAGTGTGTTTTCACAAACATGACTGCAAGGCGCTGTAAACGGGACGAT 1234
QY 2274 TTGCTTTTGGCGCTGCGGACGACGCGATCTGCGGATGCTCTCCAGGAACTCA 2333
DB 1235 TCGGATTCAGACGCTGCGGATGCAAAACGATCTCCGCGTCTTATGCAACCTGCTGC 1294
QY 2334 GCGCGCTGCGCGACCGGCGCGCTGCTGCTGTTTGCCTGATTCAGATCTGCGCAGCA 2393
DB 1295 ACCTGCGCGGCAACCACTCGAACCTCGATTTACAGATGATCAATGAATGGCAGTA 1354
QY 2394 CCACGATATGCTATGCGCGGAACTGCGCGGCTGCGCAACGCGCTGCGAGGTCGATGGCG 2453
DB 1355 CAGCGATATGCTTTGCGGGAATCTGCCGCGCTTTTCAATGCGTTGCAATGAGGGCG 1414
QY 2454 TGATTCGCAATCAATGGAAGCGGTGATGCTGTGTGCAAGAGCGTTACAGCTGCGT 2513
DB 1415 TGATGCTGATCAATGAGCGCGGAGGTGAGTATGCAAGAAAGGTCAAGGCTCCGCT 1474
QY 2514 TCGTGTGCGGTGCGCGGCTGCGGCTGCGGCTCAATGCGCAACGCGCATTCGCTGTGTGTA 2573
DB 1475 TTGTTTCGTTGCGGCTGCGGCTGCGCTCAACGCGCAACCGGGTTTTCCTTGGCGGTGA 1534
QY 2574 TGCCTTTGCTTGTGCTCAGATGAGAAGCGCTGACGCGCTTACAGGCGCAGAGTGA 2633
DB 1535 TGCCTTTGAGTACGAGCAGCAGGAGTGCAGCTTGGAAGCGCTTACACACAGGAAAAA 1594
QY 2634 TCTACAGCGCATCATGCTGCTCATGCGCTGTCAATGCGCTGTCAATGCGCGGCTTGC 2693
DB 1595 TTATATGATGCGGTGATGCGACGTCAGATCGTGTATGCGGATCATATGATGAGAAATG 1654
QY 2694 GCTGCGCGAAGCGCATGCGTTTACATGATGCTGTGCGCGCTGGCGCAATATGATCAGC 2753
DB 1655 GTTTTACCGCGGTGAAAACATGCAATGTTTCTTTCACATGGGCAAAATACGCAAGT 1714
QY 2754 TGGTGCCTGCTTTTGAATTTTTCACGCGCAGCACTGCGACGCTGCTATCAAGGTTGGGT 2813
```

```
Db 1715 TGAATCCCGAATCGATTTTCCCGCAAAACCGCTGCGCACTGCTTTTCATGCGGTTGAC 1774
QY 2814 CGCTGCGAATCTCAGTTGCTAGCGCGCGCTCGCCGCAACCTG---GCGACGCTGCGC 2870
Db 1775 CGTTACGGCAACCCCGAGGGGACCGCGGGGTCATCACTTCTTATTTTCCGTCGCCGACA 1834
QY 2871 AGCCGGGTGATATGCTCGCTGCGGCAACGCTACAGAGGGGATCGCTTTGCTGCTTTGCTG 2930
Db 1835 AACCCCGATTTTTCCTCGCTGCGGCACTCGACAGGGGACATCGTTATGCGCTTTGACA 1894
QY 2931 ATCTGCTGAGCGCTGCGGCAATCAGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2990
Db 1895 CCATCGCCAAAGCTCTGGAAGAGTGATCGCATTTACTGTTGGACACATTTGCGGCGC 1954
QY 2991 TGAACCGCAGCCAGGACATCACTGAGCTGAGCTGCGCTGCTGCTGCTGCTGCTGCTGCTG 3050
Db 1955 TCTCAGCAGCAGGCGAGGCTGGAATGCGCCGCGGGCGGGAATTGAGTTGTTG 2014
QY 3051 TGGATCAGCGCGCGCGCTGAGACATGCGCACTGTTTATCCTCAGCGCGCTGCTGACA 3110
Db 2015 CCATCATCTCCGAGCATTTCAAGGCAAGTTGACATCACATGCTGAGGATGATA 2074
QY 3111 GTGCGCTGGAAGGACTGGAAGTGTGCAAGCGGAGTGTGCTGCTGCTGCTGCTGCTGCTG 3170
Db 2075 CGGTACTGAGCGTTATTTGCTTCCCGCACCGCTACTGCGCTGCGCTGCGCATTTGATC 2134
QY 3171 AGCCCGCGCTGCGGCACTGATTTAGTGTGCAAGCGCGCTGCGCGCGCGCTCAGTTTCA 3230
Db 2135 AACCTGCGCGGATCAGCATCAAGATTTGTTATCATGGCAATCGGCAAGCGCGCTCGCTG 2194
QY 3231 GCGCGGTGCGGCACTGAGACACCACTGCAACATTTGCTGAGTGTGAGATGCTATGCTC 3290
Db 2195 CTACCAAGCATGCGCTGCGCGCGGCAATTCATGCTGCTGCTGCTGCTGCTGCTGCTG 2254
QY 3291 TGGCATGTCAGCATTTCAAGCGGAGCTGAGCGGCGCGGCTGCTGCTGCTGCTGCTGCTG 3350
Db 2255 AGCGTATGACAAATTCAGCGCGCATTTGCTGCTGCGGAGGAGCACACGCGCGCGC 2314
QY 3351 ATATTGTCAGCAGCGCTGTGTCAAGCAACAAATGCTGTGCGCGAGGCGCACTGATGCG 3410
Db 2315 ATATTGTTGAACAGGGGATGCGGACCTGTCAAGCAAGTACT-CAGTGGGAGGATTTAGCA 2373
QY 3411 CGCACTTATGATGCTATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3470
Db 2374 ACCGCACTATGATCTATTCGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2433
QY 3471 TTTAAGCGAGCTGAGCGCGCACTTAAAGTTTGTCTACTGAGAGTCAAGCGGAGCGCG 3530
Db 2434 GCTTCAGCAACAGCATTCGGATATGCGGATCTTGTATTGAGGCGGCTCTGAGGCGG 2493
QY 3531 CGGCAATCATACCTGCTGCTTCCATCGCGAAGAGCTGACGGAAGCGAGTTTCTGCTGCT 3590
Db 2494 AGGGAACATACCTGCTGCTTCCATCGCGAAGAGATTTAAAGCTGATACAGATGCTGAT 2553
QY 3591 CGAGCGCTGCTTTCGCGCGCTGCGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3650
Db 2554 AGGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2613
QY 3651 CGAGCTGAGTGTGATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3710
Db 2614 CATGTGAACAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2673
QY 3711 GGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3770
Db 2674 AAGTTTGAACAATTTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2733
QY 3771 CAGATGCGCGGATGCGGAGTGTTAACAGCGCGAGCGGCTGATTTAGCGAGCGCGGCTGCA 3830
Db 2734 CCGATTAGCGGATGCGGAGTGTTAATTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2793
QY 3831 GCGGACACCGCATCTGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3890
Db 2794 GCTGATTCGCACTAAGCGTAGGATTCAGAGCATTTATCGGTCAAGAGTGGCAACTGAG 2853
QY 3891 CGGCGCGCATGCGCTGAGAGCAATGATTTATGAGCGCGAGCTGCTGATGAGACAGG 3950
Db 2854 CGGCGCGCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2913
QY 3951 TTAATGCTTGTGTTAAGCTGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4010
Db 2914 CTACGCTTGTGTTAATCCTGCGCTTTCGCGCAACCGCATCTGCTGCTGCTGCTGCTGCTG 2973
QY 4011 CTACATCAACATGCAAGCTGATGCTGCAAGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTG 4070
Db 2974 CTACATGCAAGGCTGATCTTCAAGCGCGCAAGCGCGCGCTGCTGCTGCTGCTGCTGCTG 3033
QY 4071 CCACAGCGCGCTGGAATTTGCGCAAGCTGCTGCGGAGAGACAGGCTGCTGCTGCTG 4130
Db 3034 TGGCGCAAGAGGTGCGCTGTTACAGAGTGTGCTGCGGGAAGAACAGGCTGCTGCTG 3093
QY 4131 CAGGCTCAGCGGATGATGATGATGCTGCGAACAAGCAGCAGGCGCAACCGTGCAGCGG 4190
Db 3094 TACGTTAAGCGGATGATGATGATGCTGCGAACAAGCAAC---CGCAAGCTGTAGCGG 3150
QY 4191 GCTGCGCGCGGACTGTTTCAAGCGCACACCGGTTACTGCTGCTGCTGCTGCTGCTGCTG 4250
Db 3151 ATTAACGCGCGGCTGTTTCACTGCGACACACCGGCTACTCCTTACCGCTGCGGCTG 3210
QY 4251 GCGGAGAAATTTGCCAGACGCTGCGCGCGCAAGCTCAACAGCTGAGCCATGATCGA 4310
Db 3211 GCGCGATGCTCAGCGCGCTGATGATGTTTACTCTCTGCTGCTGCTGCTGCTGCTGCTG 3270
QY 4311 ATCTTTGCGCGGCTGAGCATGCGGCGAGAGCGCTTTTCCGCTGTTAAATGCGATGCT 4370
Db 3271 TCACTTGTCCAGCAACGTTGGAGCAACAGGAGTTTTCGATGCTGATGCTGCTGCTG 3330
QY 4371 GTTCTTGTGCGGAGCTGTAACAGCGCTGCGCGCTGATGCAAGCTTTTAAACCGCTTGA 4430
Db 3331 GTTTTTAAAGCGGACCGGCGGAGTCAAGCTGCGCTGATGCAAGCTTTTAAAGCTTACC 3390
QY 4431 CGCGGATGATTTAGCGCTTTTAAACCGGCGCACTGCGCTGAGGATTAAGCAGCGAT 4490
Db 3391 CGAGATTTGATGCTGCGCTTTTATGCGGGAATACTACCGTGAAGCAGTGGCTAGCGAT 3450
QY 4491 TCTGTGCGGCAACCGCGCGCTCTGCGGCGAAGCGCTGCGGCTGATGATGATGCTG 4550
Db 3451 TCTGACCGGCAACCGCGCGCTCTGCGGCTGCGGCTGCGGCTGATGATGATGCTGCTG 3510
QY 4551 TCGGTTACAGGGAATAATGAAACGCACTGATGATGATGATGATGATGATGATGATGATG 4610
Db 3511 TCGTTAAGAGCACTA---CATGAACCAACTAGGATTTGCTGCGGCTTTGCTGCTG 3567
QY 4611 CTGCGCTGCGCATTTGCTGCAAGCGCGCGGCTGCGGCTGCAAGCTGCTGGAACAGCG 4670
Db 3568 CTGCGCATGCGCAATTTGTTTACAGCGCGAGGATTTCTGTTTGTGCTGCTGAGCAGCG 3627
QY 4671 GATTAAGCTGCGCGCGCGCTGATGCTGATCAGGATCAGGCTTTTAACTTGAATGCGGCT 4730
Db 3628 GACAAACCGGCTGCGCGGCTTATGTTTATCAGGACAGGCGCTTACTTTTATGAGCG 3687
QY 4731 CGAGCGGTATTAACGATCCAGCGCTATCGAGGCGCTGTTTACGCTGCGGAGGCAAGCA 4790
Db 3688 CTACCGGTTATCAACGATTCAGCGCGGATTTGAAGAACTGTTTCTGCTGCGGTTAAACG 3747
QY 4791 CTCAGTATTTATGCTGACCTGATGCGGCTGACGCAATTTTATGCTGCTGCTGCGAAGAC 4850
Db 3748 CTTAAGATTTAGCTGAGGCTGTTGCGGCTGACGCGCTTTTATGCTGCTGCTGCGAGTGC 3807
QY 4851 GCGAGCAGCTGAGCTACAGCAACATCAGGCGCAGCTGAGAGAGAGATGCGACTTTT 4910
Db 3808 GCGAAGCTCTTCAATTAAGATTAACAGCAGGCGCAAGTTGAAAGCGAGATACAGAGTTT 3867
QY 4911 AATCCCGAGATGCTGCGGCTTACCGGCTTCTGAGCTATTCAGAGATGCTGCTGCTG 4970
Db 3868 AATCCCGGAGTGTGCGGCTTATCAGAGCTTCTTGAATTTCTGCTGCTGCTGCTGCTG 3927
```


QY	4971	GAGGCTATCTGAAATCGGCAACCGTACCTTTTCGCTATTTCCGCAACATGCTGCGTGGC	5030
Db	3928	GAGGGCTATCTGAAAGCTCGGCACCTGTGCCTTTTATTCGTTCAAAACATGCTTCGGGCGC	3987
QY	5031	GGGCAACAGCTGGGTCGGCTGCAGGCGCTGGCGCAGAGTCTACAGCATGCTGGCGAAATTT	5090
Db	3988	GGCGCCCAAGTTGGCAAAAGCTGCAAGCATGGCGCAGCGTTTACAGTAAAGTTCCCGGCTAC	4047
QY	5091	ATTATGATGATCATCTGCGCCAGGCTTTTTCCTTCACTCGTTGCTGGTCCGCGGTAT	5150
Db	4048	ATTGAGATGATGATCTTCGGCAGGCGCTTTTCTTTTCACTCGCTCTTAAGTGGGGGGGAAT	4107
QY	5151	CCTTTTGGCAAGCTCTTGATCTTATACCTTAAATTCAGCACTGGAGGGCGAATGGGCGAG	5210
Db	4108	CCGTTTGGCAACTGTGCTCAATTTAATACCTGATATTCACGCTTAAACCGGAATGGGCGTTC	4167
QY	5211	TGTTTTCGCGCGGCGGATACCGGTCGCGCTGGTGTGATGGCATGGCGCGGCTGTTTCGCAAT	5270
Db	4168	TGTTTTCCACCGGTTGGAACCGGTCGCGCTGGTCAATGGCATGATCAAGCTGTTTCAAGAT	4227
QY	5271	TTGGGCGGTGAACCTGCTGCTCAACGCCGAAGTCAAGCTGGAGACCGAGGGTTAACCCG	5330
Db	4228	CTGGGCGCGCGAAGTCTGTGCTTAAAGCCCGGCTCAGTATATGAAACCGTTGGGAGCAAG	4287
QY	5331	ATCAGCGGTGTCCAGCTGAAGATGGGGGCGCGTTTTCGCCCGCGCGCTTGTCGAAT	5390
Db	4288	ATTCAAGCGCGGTGCACTTGAGAGACGGCAGACGGTTTAAACCTGCGCGGTGGCGTGAAC	4347
QY	5391	GCTGACGTGTGCATACCTACGATCGCTGTTAAAGCAAGCATCTGTGGGCGGTAAACGC	5450
Db	4348	GCTGATGTGTGAACATACCTATTCGCGATCTGTGCTCTCAGACATCCGCAAGCGCTAAGCAG	4407
QY	5451	GCGGCAACGCTGGAAGCGCAAGCGGATGAGCAACTCGCGTTTGTACTTATTTTGGTCTT	5510
Db	4408	GCGAATAAACTGCAATCCAAAGCGATGTAGTAACTCACTGTTTGTACTTATTTTGGTCTC	4467
QY	5511	AATCATGCCCAACCCGCACTGGCGCACCAACGAGTGTGCTTTGTCCGCGTATCGTGA	5570
Db	4468	AACCATCATCAACGATCACTCGCCCATATACCTGTGTTTGGGCAACGCTACCGGTGA	4527
QY	5571	TTGATTCGATGAGATCTTCAATATAGCAACGAGCTGGCGGAAGTTCTCGCTGATCTGCAAT	5630
Db	4528	CTGATTCACGAAATTTTAAACCAAGATGATGCTGTGGGTGAGGATTTTGTCTTATTTTACAC	4587
QY	5631	GGCGCTGCTCAGCGATCCGTCGCTGGCACCGCGGGCTCGCGCAGTTTTCATGTCGTG	5690
Db	4588	GCACTTGTGTCAACGATCCGTCATCTGGCAACGGAAGGTCGCGCAGCTATTAATGTCTG	4647
QY	5691	GGCGCGGTGCGCATCTGGTACCGCCGCAATTTGACTTGGCAACAGAAAGGCCCGCGCTTG	5750
Db	4648	GGCGCTGTCCACACTTGAAGGCAACGCGCAACCTCGATGGGCGGTAGAAAGGCCCGCACTG	4707
QY	5751	CGCGATCGATTTTGGCTTATCTGAGAGGACACTTATGCGGGGTCTGCGACACAGTTTA	5810
Db	4708	CGGATCGATTTTGTGACTTACTTATGACAACTTACATGCTGTGGTTTCCGAAGCAAGTG	4767
QY	5811	GTTGACACACCGTATGTTTACGCGCTTTGATTTTTCGCAACGCTGACAGCGCATCAGGCG	5870
Db	4768	GTTGACACACCGTATGTTTACGCGCTTTGATTTTTCGCAACGCTGACAGCGCATCAGGCG	4827
QY	5871	TCAGCGTTTTCGCTGGAACCACTTTTGAACGAAACGCGTGGTTCCGCGCGCATTAACGC	5930
Db	4828	TCGCGCTTCTCGGTTGAACTTATCTGACCCAGAGCGCTGTGTTCCGACACATTAACCG	4887
QY	5931	GATGCGCATTACTTAACCTTTATCTGTGTGGGGGTGGCAGCATCCCGGTGCGGTGTG	5990
Db	4888	GATTAACCATTTGATTAATCTTTATCTGTGTGGCGCAGGCAACCATCTCGGCGCGGCAAT	4947
QY	5991	CCAGGCGTGTATCGGCTTCGCCGAAAGCAACCGCCAGCTGATGTGTGGAAGATCTGACCGGA	6050
Db	4948	CCGCGGTATATCGCTTCGCCGAAAGGAGGACGCGAGCTTATATGCTGAAAGACTGATTTTGA	5007

[illegible]


```

Db      6079 TAATGACGGAACGGCCAGCGCTGCTGTACCACTGCGTCTGTGACCACTTAAATAACAGTAA 6138
Qy      7191 GCCATAGGTCGTCATGCGACGCCCAATCACTGACGGCGCAGACGCCGTTGACACCCAG 7250
Db      6139 ACCATAGCGGTCATGCTGCTGACCAATCCACTGGAGGGCGCAGATTCTCTGACTGCGGAA 6198
Qy      7251 CCAATACAGGCCAATCGACACAGCGCAACACCAACCGGCTATACAGTCTGTTAGCTTAA 7310
Db      6199 GTAAATCAGGGCAATCGACCAATGGGAATACCAACGCGCATAGAGATCGTTAACTTCAA 6258
Qy      7311 TTGCTCTCATGTGCTCATGATGCGCAATGCGACGCCCATGCCAGCCATGACATGAT 7370
Db      6259 TGGCGCTTACGCGGTTTCATGATGTGAAAGATGCGACCCCAACCCAGCGTGCATGAT 6318
Qy      7371 GATTTATGCGACAGCGCCGCGAGCATCTTCATCATGATATGACGTTAGCCAGAGATAG 7430
Db      6319 GATTTATGCGACGTCAGCAACCACTTCATGCGCGACCAACGCGTGAACAACAGATCAG 6378
Qy      7431 CGCATTCATACCAAGCATCATTTGCTCCATTTGGCAAGATGAGATTTAAAGTGA 7490
Db      6379 GGCATTCCAATCCAAACATTAATCTCTCCAGTACAGACGTCGCGCATCAACGTCAGGA 6438

RESULT 3
LOCUS   ERWCR1 6918 bp DNA linear BCT 24-AUG-2002
DEFINITION Erwinia uredovora DNA carotenoid biosynthesis gene cluster (crte,
ACCESSION D90087
KEYWORDS D90087.2 GI:22474502
SOURCE   Pantoea ananatis
ORGANISM Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
          Enterobacteriaceae; Pantoea.
REFERENCE 1 (bases 1 to 6918)
AUTHORS  Misawa,N., Nakagawa,M., Kobayashi,K., Yamano,S., Izawa,Y.,
          Nakamura,K. and Harashima,K.
          Clucidation of the Erwinia uredovora carotenoid biosynthetic
          pathway by functional analysis of gene products expressed in
          Escherichia coli
          J. Bacteriol. 172 (12), 6704-6712 (1990)
JOURNAL  91072214
MEDLINE  2254247
PUBMED   2254247
COMMENT  On Aug 23, 2002 this sequence version replaced gi:216681.
          These data kindly submitted in computer readable form by: Norihiko
          Misawa
          Marine Biotechnology Institute
          3-75-1, Heita, Kamaiishi-shi,
          Iwate 026-0001
          Japan
          e-mail: norihiko.misawa@mbio.jp
          Phone: 81-193-26-6581
          Fax: 81-193-26-6584
FEATURES
     source          1..6918
                     /organism="Pantoea ananatis"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:553"
                     /note="carotenoid biosynthesis genes"
     gene            225..1133
                     /gene="crte"
     CDS             225..1133
                     /gene="crte"
                     /codon_start=1
                     /transl_table=11
                     /protein_id="BAAL1424.1"
                     /db_xref="GI:216682"
                     /translation="MTVCARKVHLTRDAEQULADIRRLDQLPEVGERDVGAM
                     REGALAKRIRPMLLTPARDGCAVSHDGLDACAVEMVAASLIIDPMCMDDA
                     KLRGRPTIHSYGEHVAITLAAVALSKAGVADGLPLAKNRAVSELSAIGMO
                     GLVQGGPKDLSBDKRSAAAILMTNHTFKSTILFCAISMQMASTIVANAASSARCLHFR

```

```

     gene            1143..2438
     CDS            1143..2438
                     /gene="crtx"
                     /gene="crtx"
                     /codon_start=1
                     /transl_table=11
                     /protein_id="BAAL1425.1"
                     /db_xref="GI:216683"
                     /translation="MSHPAALAPPFYSYRALONLQAEVARGHRYTFIQYDIKHLI
                     DSEITFHSYDSDHPGALTRYLHLAHPDGSMKLTINEMARTDMLCRELPDAR
                     DIAVDGIVDQMEPAGALVAELGLPEISVACALPINRBPDMLAVMPPEYGSOPAR
                     ERYABEKIYDMLMRHNDVYIAHSHRMGLAPQKLHQCSPILAQISQVLEIDFPK
                     ALPACFAVGPPLRETHAPSTSSRYFTSSKEPIFASLTGLQGRYGLFETIVAKEE
                     IDQULIAHGRITDQCEBLASRHTQVDFPDQSAALSOQLATHTGGMTVLVAI
                     NYRPLIALPLAFEDQGVASRIYVHIGIKASAFSTSHALARMRSILTFVDDQORWA
                     KIOTLRILAGTMAADIIEQWCTGQPVILSGSVATL"
     gene            2422..3570
     CDS            2422..3570
                     /gene="crty"
                     /gene="crty"
                     /codon_start=1
                     /transl_table=11
                     /protein_id="BAAL1426.1"
                     /db_xref="GI:216684"
                     /translation="MQPHYDLILVAGLANGLIALRLQOOQPMRILLIDAPQAGN
                     HTWSFPHNDLTESQHRMIAPLVVHMPDQVREPTTRKRLNSGYFCITISQRFABVLR
                     OFGPHLMQDTVAEVAESVRLKGGVIGAPAVIDGRYANALSGFQAFQGEWR
                     LSHPHGLSSPILIMDATVDQONGRFVSLPLSPTRLLIEDTHYIDNATLIDPCAROI
                     CVYAAQCGMQLTLREBOGALPITLSGNADATWQORPLACSLRGLFHPPTGYSLP
                     LAVAVADRSLADVFTSASIHAIHIFABERWQOQFPRMLNMLTLACGADSRMRV
                     ORRYGLPEDLIAFVAGKLTLDRLRLISGKPEVPVLAALQALMTTTR"
     gene            3582..5060
     CDS            3582..5060
                     /gene="crtI"
                     /gene="crtI"
                     /codon_start=1
                     /transl_table=11
                     /protein_id="BAAL1427.1"
                     /db_xref="GI:216685"
                     /translation="MKPTTVIGGFGGLALAIRLOAGIPVLLIEORDKRGRAYYE
                     DQGFPTDAPRYITDPSAIEELFALAGKOLKEVELLPPTPRFLRMEGSKVNVNND
                     QTRLEAQIQFNRDVEGYRQFLDYSRAVFKESYGLGGLVPLPSFEDMLBAQOLAK
                     QAWRSYVSVAYSIEDHLRQAFSPHSLVGGVGFATSSIVTLLIHLERMGVAFPRG
                     VGTALVQGMIKLFODLGGSVLVNARVSHMETGNKLEAVHLEDGRFLTQAVSNADV
                     VHTYRDLFSQHPAAVKSNNKLOTKRMSNLFVLYFGLNHHLDQLAHFTVCGFYREL
                     IDEIFNHDLAEDFSLYLHAPCVTDSLAPEGGSYYLAPVHLGTANLDMVVEGPK
                     LDRITAYILEGHTMPELSQSLVYTHRFTFPDPDQANAHGSAFSVBPVLTQSAWFRP
                     HNRDKTITNLVYVAGTHPGAGIPGVISAKATAGLMLBDLI"
     gene            5057..5986
     CDS            5057..5986
                     /gene="crtB"
                     /gene="crtB"
                     /codon_start=1
                     /transl_table=11
                     /protein_id="BAAL1428.2"
                     /db_xref="GI:22474503"
                     /translation="MNPSLNLNAVETMAVGSKFATASGLFDKTRRSVLYLWYMR
                     HCDVDVDDQTLGFOARQAPALOTPEORLIMLEKTRPAVYGSQSHNEBAPAFQVNAAM
                     DIAPAYAPHLBSGFANDVREAOYSQDDTLRTKYTHAYGVGLMAADIMKQVDAITDR
                     ACDLGLAFQULINARDIVDHAAGRCYLPASWLEHGLNKENTAAAEENQALSRITAP
                     LVQEARPYLISATAGLAGLPLRSAMAIATAKQVYRKIGVKEVQAGQADQADROSTTTP
                     EKLTLILAASQALTSRMRAPPRPRLMORPL"
     gene            complement(5925..6452)
     CDS            complement(5925..6452)
                     /gene="crtZ"
                     /gene="crtZ"
                     /codon_start=1
                     /transl_table=11
                     /protein_id="BAAL1429.1"
                     /db_xref="GI:216687"
                     /translation="MLMIMNALIVPVTYIGMEVIALAIAIKYIMHGKMGHLSHHEPR

```

ORIGIN

KGAFVNDLYAVVPAALSTIIYIGSTWMPLOWIGXMTAYGLYPMNDGLVHQRW
PFRTPKGYIKRLIMAHRMHHAVRKGECVSEFLYAPPLSKLOATLRHRHARAGA
ARDAQGEDEBPASGK"

Query Match 30.2%; Score 2618; DB 1; Length 6918;
Best Local Similarity 64.7%; Pred. No. 0; Mismatches 2165; Indels 24; Gaps 7;
Matches 4011; Conservative 0;

QY 1284 CTTGAACATTATTTATTCCTGCGGAGCAGAGCCGATCGCGTTGCGGCGCCAGCGTGC 1343
DB 303 CTTATATGATTATTCGCCGTGAGAGAGAAACGAGATGTTGTGGTGCAGCATGCGTAA 362
QY 1344 GGCACGCTGGACCGCGGAAACGTAATGCTCCGCTCTTGTCTGCTGCTGAGCAGCGAT 1403
DB 363 GGTGCGCTGGACCGCGGAAACGTAATGCTCCGATGTTGTCTGCTGAGCAGCGCGGAT 422
QY 1404 ATGGAGCTGTGACGTGGCGCAGAGGGCATCCTTGATCTTGGCTGTGCGGTGCAATGATG 1463
DB 423 CTGGGTTGCGCTGTCAAGCATGACGATTAATGATTTGGCTGTGCGGTGGAATGCTC 482
QY 1464 CACGCTGCTCACTGATCTTGCAGACATTCATCAATGATTAACGCCGAGTGCAGT 1523
DB 483 CACGCGGCTTGCCTGATCTTGAAGATGCGATGCGCATGACGATGCGAAGCTGCGCGC 542
QY 1524 GGGCGCCCGGCAATCCACTGTGAATATGGGAAAAAGTGGCATCTTGGCAGCGGTGCG 1583
DB 543 GGAAGCCCTACCATTCATTCATTAACGAGAGCATGTGGCAATATCGGCGCGGTGGCC 602
QY 1584 CTACTCAGCCGCGCTTGAAGTATGCTCTGCGCGCGGCTGTGCCAGCAACGACAAA 1643
DB 603 TTGCTGATTAAGCTTTTGGGCTAATTTGCCATGCAATGCGCTCACGCCCTGGCAAAA 662
QY 1644 GCCGAAGCATTGCAGCTCTCTCTGCGGTGGCGCTGCAAGGACTGTCTCAGGGTCA 1703
DB 663 AATCGGCGGCTTTCGAACTGTCAAAACGCATCGGCATGCAAGATGTGTTCAAGGTGAG 722
QY 1704 TTCCAGATCTGCATGACGCGCACACAGCGCGAGTCGCGAAGCATCAACCTGACCAAT 1763
DB 723 TTCAAGATCTGTCTGAAGGAGATAGCGCGCAGCGCTGAAGCTATTTGATGACGAT 782
QY 1764 GAACGTAAACACAGCGCTGTCTGTTCGCGCACAGCTGCAGATGCGCGCAT 1823
DB 783 CACTTAAACACAGCGCTGTCTGTTCGCGCACAGATGCGCATGCGCATGTTGCGCAT 842
QY 1824 GCGTACGTGCAAGTACGTACGCTTTAAGTATTTTGGCGAGATTAAGTCAAGCTTTC 1883
DB 843 GCGTCCAGCGAGCGCGTGTGCTGCTGATGCTGCTTTCATCTTGATCTTGTGACGCAATT 902
QY 1884 CAGTTACGTGAGCGAGCTGCGGATGCGCTTAAGCACACGCGCAAGAGCATGTCATCAGAT 1943
DB 903 CACTGCTGAGCGATTTACCGATGCGATGACCGACCGGTAAAGATAGCATACAGGAC 962
QY 1944 CAGGCAATATCCAGCTGTGTCAGATGCTGGGCCCGAAGGGGCTGAGCGTGTCTGCGC 2003
DB 963 GCCGTAATAGAGCGCTGTGTAATCTGTAGGCCCGAAGGGGCTTGAAGAACGTCAGGA 1022
QY 2004 GACCATCTAAGAGCGCGGATGACACCTTGTCTGCGCTGCGCATGCGGTGTGCGCAC 2063
DB 1023 CAACATCTTCAAGCTTGCAGTGAAGCATCTCTGCGGCTGCGCAACAGGGCAGCGCACT 1082
QY 2064 CGTCAATATGACAGCGCTGTCTTAAATCAACAGCTGGGATGTTCAATGGAAGCGGCA 2123
DB 1083 CAACATTTATTTCAAGCGCTGTCTTAAACAAAACCTGCGCGCTACGTTAAAGATGCT-- 1140
QY 2124 TACCTATGGGCGATTTTCCGTTATTTGGCGCACCGCTCTACAGCCATTTTACGCGATTGC 2183
DB 1141 ---GATATAGCATTTTGGCGGAGATGCGACCGCTTTTAAAGCGCATGTTGCGCGATTAC 1197
QY 2184 AGGCGCTGGCGAAACGCTGCTGGCGCGCGGACATGCACTTATTCAGCAAAATG 2243
DB 1198 AGAATCTGCTCAGAACTGGTGGCGCGCGGTCATCGGATGACCTTTATTCAGCAATACG 1257

QY 2244 ATGACAGCACCTTGTGAGCGACGAGCGCATTTGCTTGTGGCGGTGCGGAGCGCACGC 2303
DB 1258 ATATTAACACTGATTCATAGGAAACCATTTGATTTATTCCTGTCGGACAGACAGCC 1317
QY 2304 ATCTGCGGATGCTCTTCCAGCGAACTCAGGCGGTGCGCGACCGGCGGCTGTGCG 2363
DB 1318 ATCCCCCGCGGTTAACCGCGGTGCTACACTGCGGCGGTCACTCTGTTGGGCGCGTCAA 1377
QY 2364 TGTTCGCGGATTCACGATCTGCGCACGACCGATTAAGCATGCGGAACTGCGCG 2423
DB 1378 TGTGAAGCTCATCAATGAATATGGCGGACCAACCATATGCTGTGCGGAACTCCCC 1437
QY 2424 CGGTGTCACAGGCTGACAGTGCATGATGCGTGTGATGCGTATCCATTAAGAAAGCGGTG 2483
DB 1438 AGGATTTAACGATCTGGCGGTGATGGCTCATTTGTATTAATGAATGAACGCGAGGCG 1497
QY 2484 GTCTGTGGCAGAGCGTTACAGCTGCGGTTGCTGTGGTGGTGGCTGCGCGCTGCGCTCA 2543
DB 1498 CGCTGTTGTGAAGCATGGGACTGCGGCTTATCTGTGCGCTGCGCGCTGCTCTCA 1557
QY 2544 ATGCGAAGGCGCATTCGCGGTGTGATGCGCTTGTGCTTGTGCTCAGGATGAGAAAG 2603
DB 1558 ATGTAACCGGATATGCTCCGTGGGTATGCTTTCGATACGGAACGAGCGAGCGG 1617
QY 2604 CGCTGACGCGTATCAGGCGCAGAGTACATCTACGACCGCATATGCTGATGCG 2663
DB 1618 CTGCGCAAGCTTATGCGCGCAGTAAATTTATGATGCTGCTAATGCGTGTGATGACC 1677
QY 2664 CTGTGATGCTGCTGATGCGCGGCTTGGCGCTGCGCGAAGCATGCTTACATCAT 2723
DB 1678 GTGTGATGCTGCGAACAGCGCACAGATGCGCTTGAAGCCCCCGCAAAAGCTTACACAGT 1737
QY 2724 GTCTGCGCGCGTGGCGCAATGATGATGATGCGCGCTTGTGATTTTCAAGCGCAGC 2783
DB 1738 GTTTTTCGCACTGGCGCAATGACGAGCTTGTCTGTAATGATTTTCCCGCAAG 1797
QY 2784 AACTGCAGCGCTGATCACAGCGTGGTCCGCTGCG--GACTCCAGTTGTACGCGCG 2840
DB 1798 GCTTACCGGCTTGTTCATGCGCGTGGCGCTGCGCGCAAGCAGCACCGTCAAGT 1857
QY 2841 CGCTGCGCGACCTTGGCGCGCGTGGCGCGAGCGGTGTATGCTTGGTGGCGCGC 2900
DB 1858 CTTCATCCGTTATTTTAACTCTCAGAAAAACCCCGGATTTTGTGCTGCGCGCGC 1917
QY 2901 TACAGGGCAATGCGCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2960
DB 1918 TTCAAGGACACCGTTATGAGGCTGTTAAACGATAGGAAGCTGTGAAGAAATTAAG 1977
QY 2961 TGTGCTGTGTGTGACACTGTGTGGCGGTGAACCGCACCGCAGGACATCACTCAAG 3020
DB 1978 GTCAAGTCTGTGTAGCCCACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2037
QY 3021 TGGCGGTGCTGCGGTGTGACGATTTTGTGTATGACGCGGCGGCGCTGACGATGCGC 3080
DB 2038 GAAACGCTATACACAGGTGTGTGATTTTGCATCACTGACGCGCTGCTCAAGCGC 2097
QY 3081 AACTGTTATCACTCACGCGGCTGTAACAGTGTGCTGTAAGCACTGGAATGTGGCAGC 3140
DB 2098 AGCTGGGATACCAACGCGGCGCATGATATGAGTACGACGCGATTAATTAACGAGACG 2157
QY 3141 CGATGTGGCGCTGCGCATGCGCTTGCATGACCGCGCGGTGGCGGCGCATTTGAATGATG 3200
DB 2158 CCGTTTAAAGCGCTTCCGCTGCGCTTGTATCAGCCCGCGTGTGCGCATCGTTATTC 2217
QY 3201 AGCGGCTGCGCGCGCGCTGACGTTTCAAGCGGCTGCGGACGCTGGAACACCACTGCG 3260
DB 2218 ACAGCATGCGCAACGCGCTTCCGCTTACACAGCCATGCTTTGGCTGTGATGATG 2277
QY 3261 AACAGTTGTGAGTACGATGCTATGCTGTGCGCATGACGATTTAGGCGCGAGCTGCG 3320
DB 2278 GTTATTTGTACCAAGGTGATGCTTTACAGACGCGCAAGGAAATCAAGACAGCCCTTC 2337
QY 3321 AGCGGCGGCTGTATACGCGCGCGGTATATTTGTGACAGAGCGCTGTGTCAAGAC 3380

Db 2338 GTTTGGAGGGGACACATGCGCGCTGCGATATCATAGACAGATTATGTG-CACCGGT 2396
QY 3381 AAATCGCTGCGGAGGCCACTGATGCGCGACCTTATGATGTCATTTGCTGGGTGC 3440
Db 2397 CACCGCTGCTTAATGTTGGAGCGCGCTATGCAACCGCATTTATGATCTGATTCCTGGGGGC 2456
QY 3441 CGGCGTGGCTAAGGGGTGATTTGGCTGCTTAAACGCACTGACAGCCGACCTTAAGT 3500
Db 2457 TGAACCTGCGAATGCTTATACGCTGCGCTTACAGACAGCAACCTATATGCGTAT 2516
QY 3501 TTGTCTACTGAGAGTCAGCGCGAGCGCGCAATCATACCTGCTGTTTCAACGCGA 3560
Db 2517 TTGTCTTATGAGCGCGCACCCAGCGGGGCGGGAATACATGCTGCTATTTACACGCA 2576
QY 3561 AGAGCTGAGGAAGCGCATGTTGCTGCTGACGCGCTGCTTTGCGCGCGTGGCGCG 3620
Db 2577 TGAATTTGACTGAGAGCGCAATGCTGTTGATAGCTCCGCTGTTGATCATCATGCGCGA 2636
QY 3621 TTATCAGTATGCTTCCCACTGCGCTGCGCAGCTGGATGATGTAATTTGCTCATTC 3680
Db 2637 CTATCAGATGCTGCTTCCCACTGCGCTGCTGTAAGCTGAAACAGCGGCTACTTTGTATAC 2696
QY 3681 CTGCGAGATTTTGGCGGCACTTACAGACAGTGTGCTGCGCGCTACGCAACCGCAGC 3740
Db 2697 TTCTCAGCGCTTTCGCTGAGGTTTACAGCAGACAGTTTGCCCGCATTTGTGATGATAC 2756
QY 3741 GCGGCTCAGGAGGTCTACCCACCGGAGTCAGACTGCGGATGGCGGATGTTACAGGC 3800
Db 2757 CGCGGTGCGAGAGTTAATCGGAATCTGTTGTTGAAAAGGATCAGGTTATCGGTGC 2816
QY 3801 GCAGGGGTGATTTGACGAGACGCGGCTGACGCGCAACCGCAATCTGCACTCGCTATCA 3860
Db 2817 CGCGGTGATTTGACGAGGCGGAGTTATGCGCAATTTACAGCTGCGCTGCTTCCA 2876
QY 3861 GGCATTTGTGCTCAGAGTGGCACTGCGCGCGCATGGCTGACAGGCAATAT 3920
Db 2877 GGGGTTATTTGGCGAAGATGCGATTTAGCACCGCATGTTTATCGTCTCCATTA 2936
QY 3921 GATGACGCGGAGGTGATCAGACAGAGGTTATCGCTTTGTTTACACCTGCGCTCAG 3980
Db 2937 CATGATGCGACGCTGATCAGCAAAATGTTATCGTTGATCAGCTGCGCTC 2996
QY 3981 TGGCAGCGCTTACTGATTAAGATACCACTACATCAACATGCCACGCTGATGCGC 4040
Db 2997 GCGGACCAATGTTATTTGAAAGACGCACTATATTTGATATGCGCATTAATCTGA 3056
QY 4041 ACAGGCGCGCGTCACTTACGATTAATGCCACACGCGCGGCTGGAATTTGGCGCAGT 4100
Db 3057 ATGCGCGCGCAAAATTTTGGCATATGCGCGCAACAGGTTGGAGCTTACAGACAT 3116
QY 4101 GCTGCGAGAGACAGCTGCTGCGCATCAGCTCAGCGCGATATGCAATGATTTCTG 4160
Db 3117 GCTGCGAAGAAACAGGGGCTTTACCATTTACTGTGCGGGAATGCGACCATTTCTG 3176
QY 4161 GCAACGACAGACGCGCAACCTGACAGCGGCTGCGCGCTGCACTGTTTCAACCCACAC 4220
Db 3177 G---CAGCAGCGCGCTGCTGATGATTAACGTCGCTGTTTCACTCAAC 3233
QY 4221 CGGTTACTGCTGCGCGCGGTTGGCGTGGCGGGAAGATTGCCAGACGCTGCGCGC 4280
Db 3234 CGGCTATTTACCTGCGCTGCGGTGCGTGGCGACCGCTAGTGCACTTATGTTT 3293
QY 4281 CGACGCTCAACGCTAGCAGCATGCAATCTTTGCGCGTACAGCACTGGCGCGACA 4340
Db 3294 TAGCTGGGCTCAATTCACATGCAATTCAGATTTTGCCGGAAGGCTGGAGAGACA 3353
QY 4341 GCGCTTTTTCGCTGTTAATGCAATGCTGTTTCTTGCGGACGCGCTGAACAGGCTG 4400
Db 3354 GGGCTTTTTCGCTGATGCTGATGCTGTTTGAAGCGGACCGCGCATTCAGGCTG 3413
QY 4401 GCGGTTATGACAGGTTTATACGCGCTGACGCGGATTTGATAGCGCTTTTACCGCG 4460

Db 3414 GCGGTTATGACAGGTTTATGCTTATACCTGAAGATTTAATGCGCGTTTATGCGG 3473
QY 4461 GCACTGCGCTTAGGATTAAGACAGCATTTCTGTGCGCAACCGCGGTCCCTCTCGG 4520
Db 3474 AAAACTCAGCTACCGATCGGCTACGTAATTTGACAGCGGCAACCGCTGTTCCGTAT 3533
QY 4521 CGAAGGCTGCGGCAATGATGATGACCTCTCGGTTACGAGGGAAGAAATGAACGC 4580
Db 3534 AGCAGCATTTGCAAGCAATATGACACTCATCTTAAAGGCACTA---CATGAACCA 3590
QY 4581 ACCTATGATTTGCGCAGGCTTCCGTGCGCTGCGCTGCGATTTGCTGCAAGCGGC 4640
Db 3591 ACTAGGTAATTTGCTCAGGCTTCCGTGCGCTGCGCACTGCGCAATTTGCTACAGGCTG 3650
QY 4641 GCGCTGCGGCTCAGCTGCTGGAACAGCGGATTAAGCTGCGCGGCGCTATGCTAT 4700
Db 3651 GGGATCCCGCTTACTGCTTGAACAACGATGAACCCGCGCGTGGGCTTATGCTAC 3710
QY 4701 CAGGATCAGGTTTACCTTTGATGCGCGGTCGACGCTGATTAACGATCCAGCGCTATC 4760
Db 3711 GAGGATCAGGTTTACCTTTGATGAGCGCGCACTGATTAACGATCCAGTGCAT 3770
QY 4761 GAGCGCTGTTTACGCTGCGAGGCAACCACTCAGTATATGCTGACCTGATGCGGTG 4820
Db 3771 GAGGATCTGTTTACGCTGCGAGGAAACAGTTAAAGATATGCTGACCTGCTGCTG 3830
QY 4821 ACGCCATTTTATGCTGCTGAGGAAAGACGCGGCACTGATTAACGCAACATCAG 4880
Db 3831 ACGCCGTTTATCCGCTGCTGAGTCAAGGAAAGTCTTAAATTAACGATACATCA 3890
QY 4881 GCGGCTGAGAGAGAGATGCTGCTTAAATCCCAAGATGCTGCGGCTTACCGCCAG 4940
Db 3891 ACCGCTGAGAGCGGATTCAGAGTTAATCCCGATGCTCAAGGTTATGCTCAG 3950
QY 4941 TTCTGAGCTTATTCAGAGATGTTTCCGAGAGGCTATCTGAACCTGAGCAACCTACT 5000
Db 3951 TTCTGAGCTTATTCAGAGCTGCTGTTTAAAGAGCTATTAAGCTGCTGCTGCT 4010
QY 5001 TTCTGCTATTTCCGAGCATGCTGCTGCGCGGCAACAGCTGCTGCTGAGGCTGCG 5060
Db 4011 TTTTATGCTTACAGAGATGCTTCCGCGCGCACTCAACTGCGAATGCGAGCATG 4070
QY 5061 CGAGTGTCAACAGATGCTGCGGAAATTTATCATGAGATATCTGCGCGCATTT 5120
Db 4071 AGAAGGTTTACAGTAAAGTGTGCACTATCATGAAGATGAATCTGCGCGCGGCTT 4130
QY 5121 TCCTTACCTGCTGCTGCGCGGCTTATCTTTGCAACGCTTCACTATCACTTA 5180
Db 4131 TCCTTACCTGCTGCTGCTGCGCGGCTTATCTTTGCAACGCTTCACTATCACTTA 4190
QY 5181 ATTCAAGCACTGAGCGCAATGCGGCTGCTGCTTCCGCGCGGCTGACCGGTGCGCTG 5240
Db 4191 ATACAGCGGCTGAGGCTGAGTGGGCGCTGCTGCTTCCGCGGCGGCAACCGCGCATTA 4250
QY 5241 GTTATGAGCATGCGCGGCTGCTTCCGCAATTTGGCGGTGAACTGCTGCTCAACCGGAA 5300
Db 4251 GTTCAGGAGATGATTAAGCTGTTTCAAGATCTGGGGGGAAGTCTGTTTAAACGCAAG 4310
QY 5301 GTGAGCAGCTGAGACGAGAGGTTAACGATCAGCGGTGCTCAGCTGAAGATGAGGCG 5360
Db 4311 GTGAGCAGCTGAGACGAGAGGTTAACGATCAGCGGTGCTCAGCTGAAGATGAGGCG 4370
QY 5361 CGTTTTCGCGCGCGCTGCTGCTCAAAATGCTGACGTTGATCACTTACATTCGCTG 5420
Db 4371 AGGTTCTCAGAGAGCGCTGCGCTGCAATGCAAGATGCTTCACTATGCGACCTG 4430
QY 5421 TTAAGCAGATCTGCGCGCGCTTAAACGCGGCAACCTGAAGCGCAAGGAGATGAGC 5480
Db 4431 TTAAGCAGACCTGCGCGCGCTTAAAGCAGTCCAACTCAAGCTGAAGGCGCATGAGT 4490
QY 5481 AACTGCTGTTTGAATCTAATTTGCTTAAATCATGCCACCGCGAGCTGCGGCAAC 5540
Db 4491 AACTCTGTTTGTGCTTATTTGTTGTTGAATCAACATCATGATGATGCTGCGGCAAC 4550

QY 5541 ACCGTTGCTTTGGTCCCGCTATCGTGAATTGATGATGAGATCTTCAATAGACGACG 5600
 DB 4551 ACCGTTGTTTGGCCCGCTACCGGAGCTGATTAACAAATTTTAAATCATGATGAC 4610
 QY 5601 CTGGCGGAAATTTCTGCTGATATCTGCATGCGCCCTGCTCCAGAGATCCGTCCGTGGCA 5660
 DB 4611 CTGGCAAGAGACTTCTACCTTATCTGACGCGCCCTGTGTGACAGGATTCGTCTGCGC 4670
 QY 5661 CCGGCGGCTGCGGAGTTTTCATGTCGTGCGCGCTGCGGATCTCGGATCCGCGCA 5720
 DB 4671 CTTGAAGTTGCGGAGTACTATGTTGTGGCGCGGCTGCGCATTTAGCGACCGCAAC 4730
 QY 5721 ATTGACTGGCAACAGAAAGGCGCGCTGTGCGCATGCGATTTTGTCTTATGAGAGAG 5780
 DB 4731 CTGACTGGACGTTGAGGGGCAAAACTAGCGACGATTTTGTGGTACCTTGACGACG 4790
 QY 5781 CACTATATGCGCGGCTGCGACAGCAGATTAGTGAACACCGTATGTTTACGCGCTTGTAT 5840
 DB 4791 CATTACATGCTGCTTACGAGAGTCAAGCTGTGACGACCGGATGTTTACGCGCTTGTAT 4850
 QY 5841 TTTTGGCAACAGCTGACAGCGCATCAGGGCTCAGCGTTTCTGCTGAAACCCATTTTGAAG 5900
 DB 4851 TTTGCGACACGCTTAATGCTTATCATGCTCAGCGTTTCTGTGAGCGCGTCTTAC 4910
 QY 5901 CAAAGCGCTGCTTCCGCGCGCATTAACCGGATGCGCATTAACCTTATCTGATG 5960
 DB 4911 CAGAGCGCTGCTTCCGCGCGCATTAACCGGATGCGCATTAACCTTATCTGATG 4970
 QY 5961 GGGGCTGACGATCCCGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGG 6020
 DB 4971 GGGGCTGACGATCCCGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGG 5030
 QY 6021 GCCGAGTGTATGAGAGATCTGACCGGATGAACCAACCGCGCTGATGAGAGTCA 6080
 DB 5031 GCAGTTGATGCTGAGAGATCTGA---TATGAATTAATCCGTCCTTACTCAATCATGCGG 5087
 QY 6081 CGCAACCATGCGCGCGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGG 6140
 DB 5088 TCGAAGCATGCGCGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGG 5147
 QY 6141 CAACGCGCGCGCATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATG 6200
 DB 5148 AAACCGCGCGCGCATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATG 5207
 QY 6201 ATGCTCAGACGCTGCGGAGAGCGGACGACGACGCGGCTGAGAGAGAGAGAGAGAGAGAGAG 6260
 DB 5208 ACAGATCAGACGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGG 5267
 QY 6261 TGCGGCGCGCGCATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTG 6320
 DB 5268 TGATGCAATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 5327
 QY 6321 CGTTTCGCTTTCAGAGAGTGGGCTGACGATCAGCTTCCCGACGAGCTGCTTTT 6380
 DB 5328 CGTTTCGCTTTCAGAGAGTGGGCTGATGCTGATGATGATGATGATGATGATGATGATGATGATGATG 5387
 QY 6381 ATCATCTGGAAGGTTTGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 6440
 DB 5388 ATCATCTGGAAGGTTTGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 5447
 QY 6441 CGCTGCGCTTACGCTATCAGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATG 6500
 DB 5448 CGCTGCGCTTACGCTATCAGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATG 5507
 QY 6501 GCGTACGCTATGAGCGCTGATCAGCTGATCAGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATG 6560
 DB 5508 GCGTACGCTATGAGCGCTGATCAGCTGATCAGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATG 5567
 QY 6561 CCAATATGCAAGGATATGCTTGAAGAGCGGAGAGATGCGCTTGTCTATCTGCCACAA 6620
 DB 5568 CCAATATGCTGCGATATGTTGAGAGATGCGGATGCGGCGCTGTTATCTGCCGCGCA 5627

QY 6621 GCTGCTGATGAGCGCGGACTGAGCGCGCCGAGCTTGGCCATGCCCAATCGCGCAG 6680
 DB 5628 GCTGCTGAGAGCTGAGAGGCTTGAACAAAGAGATTATGCGGACCTGAAACCGTGACG 5687
 QY 6681 CGCTGCGCGCTGAGAGCGGCTGCTGCTGCGGAGCGCGGAGCGCTGATCACTGACGCGC 6740
 DB 5688 CGCTGAGCGCTATGCGCCCGCTGTTGGTGACAGAAACGAACTTACTATTTGCTGCCA 5747
 QY 6741 GCAGCGGCTGCGGAGATTGCGCGCTGCTGCTGCGGAGCGGAGATGCGCAACGCGCGCGC 6800
 DB 5748 CAGCGCGCTGCGGAGGTTGCGCTGCTGCTGCGGAGCGGAGATGCGCAACGCGCGCAG 5807
 QY 6801 TTTACCGGAAATTTGCGCTGAAAGTGCAGATGCGCGTCCGCGGAGATGAGATGAGCGCC 6860
 DB 5808 TTTACCGGAAATTTGCGCTGAAAGTGAACAGCGCGGCTGAGCAAGCTGCGGATCAGCGGC 5867
 QY 6861 AGCGCACAGTAAAGCGGAAAGCTGCGCTGCTGCTGAAAGTTCGCGCGCTTGA 6920
 DB 5868 AGTCAACGACACGCGCGGAAATTAACGCTGCTGCGCGCGCTGCTGCTGCGCGCTTGA 5927
 QY 6921 CTTTCCGCGCTTGTCTATCCCGAGGCGGCTGCGCGCTGCTGCGCGCTGCTGCGCGCTTGA 6980
 DB 5928 CTTTCCGCGCTTGTCTATCCCGAGGCGGCTGCGCGCTGCTGCGCGCTGCTGCGCGCTTGA 5981
 QY 5981 ACAGCGCCATGCGCTGCGGAGCGGCTGCGCGCTGCGCGCTGCGCGCTGCGCGCTTGA 7040
 DB 5982 TCTAGCGCGCTATGCTTCCCGAGGCGGCTGCGCGCTGCGCGCTGCGCGCTTGA 6041
 QY 7041 AAAGCCAAAGAAACGCAAGCTTCCGCTCCCGGACCGCATGATGATGATGATGATGATGATGATGATG 7100
 DB 6042 GAAAGCCAAAGAAACGCAAGCTTCCGCTCCCGGACCGCATGATGATGATGATGATGATGATGATGATG 6101
 QY 7101 ATTCAAACGCTTGAATGAGCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 7160
 DB 6102 ATTCAAACGCTTGTGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 6161
 QY 7161 CAGGCGCTGATGACCATTAATGAGAGCGCATGAGTGTGATGATGATGATGATGATGATGATGATGATG 7220
 DB 6162 CAGGCGCTGATGACCATTAATGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATG 6221
 QY 7221 CTGCGAGCGCGGACGCGCTTGTGACATCCGACCAATGAGCCAAATGCAACAGCGCAAA 7280
 DB 6222 CTGCGAGCGCGGACGCGCTTGTGACATCCGACCAATGAGCCAAATGCAATGATGATGATGATGATG 6281
 QY 7281 CACGAGCGCATGACGCGCTTGTGACATCCGACCAATGAGCCAAATGCAATGATGATGATGATGATGATG 7340
 DB 6282 AACGAGCGCATGACGCGCTTGTGACATCCGACCAATGAGCCAAATGCAATGATGATGATGATGATGATG 6341
 QY 7341 ATGCGAGCGCGCATGACGCGCTTGTGACATCCGACCAATGAGCCAAATGCAATGATGATGATGATGATG 7400
 DB 6342 ATGCGAGCGCGCATGACGCGCTTGTGACATCCGACCAATGAGCCAAATGCAATGATGATGATGATGATG 6401
 QY 7401 CATGATGATGACGCGTGTGACGAGATGAGCGATTTCCATTAACAGAGCATGATGATGATGATGATGATG 7460
 DB 6402 CATGCGCATGACGCGTGTGACGAGATGAGCGATTTCCATTAACAGAGCATGATGATGATGATGATGATG 6461
 QY 7461 ATTGCGAAGAGTGAAGTA 7480
 DB 6462 GGTGAGAGCGTGTGCGAGCA 6481

RESULT 4
 112910
 LOCUS 112910 6918 bp DNA linear PAT 26-JUL-1995
 DEFINITION Sequence 13 from patent US 5429939.
 ACCESSION 112910
 VERSION 112910.1 GI:910887
 KEYWORDS
 SOURCE
 ORGANISM
 Unclassified.
 REFERENCE
 1 (bases 1 to 6918)
 AUTHORS Misawa,N., Kobayashi,K., Nakamura,K. and Yamano,S.

D	2278	GTTCATTGCTGACCAACGTCCACTTTGACGACGCGATGGCGAAATCCAGACAGCCTTC	2337
Q	3321	AGCGGGCCGGTGGCTGTACGGCGCGCTGATATGTGAGCAGGGCGCTGTGTGACAGC	3380
D	2338	GTTTGGCAGGGGGCCAAATGGCGCTGGCGAATCATTTAGACAGGTTATATGG-CACCGGT	2396
Q	3381	AAATGATCTGGCGGAGGCACTGATGCGGGCACTTATATGTCATTTGCTGGTCGGGCG	3440
D	2397	CAGCTGTCTTAAATGGGAGCGGCTATGCAACCGATTAATATCTGATTTCTTCGGGGGGC	2456
Q	3441	CGGCTGGCTTAA CGGCGTGAATTTGCGTGCCTGTTTACGCCAGCTGCAGCCCGCACTTAAAGT	3500
D	2457	TGGAATCCCGCAATGGCCTTATTCGCTCGCTCTTTCAGACAGACGCAACTGATATGCGAT	2516
Q	3501	TTTGTACTGAGAGATCAGGGCGACCGGGCGCGCAATATATCTGTCGTCTTCATCGGA	3560
D	2517	TTTGTCTTATCGAGCGCGCAACCCCGAGCGGGGAGATATACGTGTATTTACACAGCA	2576
Q	3561	AGACCTCAGCAGAGCGGCAATTTGCTGGCTCGAGCGCTGCTTTTCGCGCGCTGGCCGG	3620
D	2577	TGATTTGACTGAGACCAACATCGTTGGATAGCTCGCTGGTGGTTATCATCTGGCCGA	2636
Q	3621	TTATTAAGTAGCTTCCCAACCTTCGTTCGCCAGCTGGAATGTGAATTTGCTGATTC	3680
D	2637	CTATCAGGTAGCGTTTCCCAACGCGCGTGAAGCTGAACAGCGCTACTTTTGTATTAC	2696
Q	3681	CTCGAGAGATTTTGGCCGGCACTTACACAGAGTGTGCTGTCGCGCGCTACACACGCGAGC	3740
D	2697	TTCTCAGCGTTTCGTGAGGTTTACACGACAGATTTGGCCGACTTGTGGATGCAATC	2756
Q	3741	GCGGTACGAGGTCACCCACCGGGGTCACTGGCGGATGGCGGAGATTACAGCC	3800
D	2757	CGCGGTCCAGAGGTTAATGCGGATCTGTTTCGTTGAAAAAGGTCAAGTTATCGGTGC	2816
Q	3801	GCAGCGGATGATTAACGAGCGGGCTGACGCCAACCGGATCGACGCTGGCTATACA	3860
D	2817	CGCGCGGTGATTAACGCGGGGTTATGCGGCAATTCAGCACTGACGCTGGGCTTCCA	2876
Q	3861	GGCATTTGTGCGTACAGAGTGGCACTGGCGCGCGGATGAGCTGCAGCAGCCAAATTT	3920
D	2877	GAGCTTTATTTGGCCAGGAATGGCGATTGAGCAACCGGATGATTTATGTCTCCATATAT	2936
Q	3921	GATGAGCCCGACCGTTCATACGACAGAGTTATGCTTTTTCATACCTTCGCGCTCAG	3980
D	2937	CATGATCCACGTCGTACAGCAAAATGGTTATGCTTGTGTGACAGCTTCGCGCTTC	2996
Q	3981	TGCCAGCGTTTACTGATTGAAGATACCCATCACTCAACATGACGACGCTGATGCGCG	4040
D	2997	GCCGACCAAGATTGTTAATTGAAGACGCCATATATTAATGAGCATTAAGATCTCGA	3056
Q	4041	ACAGGCGCGCTCATATTACGATTAATGCCAACGACGCGCTGGAAATTTGCGCCAGCT	4100
D	3057	ATGCGCGGGGCAAAATATTTTCGACTAATGCGCGCAACGAGGTTGGCAGCTTCAGACACT	3116
Q	4101	GCTGGCCAGGAGGCA CGGCTCGTCGCCGATACAGCTGACGGCGATATCATGACTTTGTG	4160
D	3117	GCTGGAGAAAGACGGGCGCTTACCATTACTCTGTGCGGCAATGCCAGCATTTCTG	3176
Q	4161	GCAACAGCAGACGGGCAACCGTACAGCGGGCTGGCGCGGACTGTTTACCGCACCAAC	4220
D	3177	G---CAGCAGCGCCCCCTGGCTGTAGAGATTAAGTTCGCTCGGTCTCATCTCAAC	3233
Q	4221	CGGTTACTCGTCGCCCGCGGTGGCGCTGCGCGGAGAAATTCGACAGACGCTGCCCG	4280
D	3234	CGGCTATTACCTGCGCTGGCGGTGGCGGTGACCGCGCTGAGTGCACTTGATGTCTT	3293
Q	4281	CGACGCTCACAGCTGACCACTGATCGAATCTTTTGGCCCGTCAAGCACTGGCGCGAGCA	4340
D	3294	TACGTCGCGCTCAATTCACATGACCATTAAGCAATTTTGGCCGAGGCGCTGGCAGCAGA	3353
Q	4341	GCGCTTTTCCGTCTGTAAATTCGATGCTGTTCTTTCGCGAGCGGCTGAAACGCGTG	4400
D	3354	GGGCTTTTCCGACGCTGAATGCGATGCTGTTTTCGCGGACCGCGCATTCACGCTG	3413

[illegible]

SOURCE) Pantoea agglomerans
ORGANISM Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
REFERENCE 1 (bases 1101 to 1300; 4301 to 4410; 4641 to 4750; 4881 to 5000;
5031 to 5140; 5561 to 5660; 6031 to 6160)
AUTHORS Liu, S.T.
TITLE Carotenoid-biosynthesis genes as a genetic marker for the purpose
of gene cloning
JOURNAL Biochem. Biophys. Res. Commun. 195 (1), 259-263 (1993)
MEDLINE 8395826
PUBMED 93371414
REFERENCE 2 (bases 1 to 6965)
AUTHORS To, K.Y., Lai, E.M., Lee, L.Y., Lin, T.P., Hung, C.H., Chen, C.L.,
Chang, Y.S., and Liu, S.T.
TITLE Analysis of the gene cluster encoding carotenoid biosynthesis in
Erwinia herbicola H9013
JOURNAL Microbiology 140 (Pt 2), 331-339 (1994)
MEDLINE 94236237
PUBMED 8180698
FEATURES
source
1..6965
Location/Qualifiers
/organism="Pantoea agglomerans"
/mol_type="genomic DNA"
/db_xref="taxon:549"
776..6687
/product="CrtE"
926..1834
/gene="crtE"
926..1834
/gene="crtE"
/codon_start=1
/transl_table=1
/product="CrtE"
/protein_id="AAA21260.1"
/db_xref="GI:148394"
/translation="MTVCAGKGVHPTSSAAEQLLADIDRLLEQLPYEGERDFVGAAM
REGALAPGKRIRPMILITLARDLGCASHEGLDIACAVENVAAHSLIDMPCMDA
QIVGGRPTVHCQYGEHVAITLAAVALLSAPFVINDAGCTLAKRAVSEISNAIGMO
GLVGGQFDLSEGGRRSAEALIMTNEFKITLTCASQMASTVANNSSERRDYLHP
SLDGAQFQLDLDITDGMADYKQINDQEGKSTLVNLSRAVERLRHLHLSSEHL
SAACQGHSTQOIFQAMFQKLAAS"
1844..3139
/gene="crtX"
1844..3139
/gene="crtX"
/codon_start=1
/transl_table=1
/product="CrtX"
/protein_id="AAA21261.1"
/db_xref="GI:148395"
/translation="MSHPAATAPPYSHVRLLENLAQELVARGHGVTFIQOQDKPLI
QPHIGFAVGEESHPGSLRLHLAHLPLSPNMLRINEMARTSDMLCIALPATN
DLADIGVLDQMEPAGALVABALGLPVSVALPLNRPDLPLAVWFDTGDLAAR
ERYASEKIVMLMRHRDRLVLAHSHHSGMLAPRDLHCFCPLAQISLIPDLPPK
ALPCFHAVPLRETQSTSTSPLEFSEKPRJFASGTLQGRHYGFKAVIACVEE
IDGQILLAHCGRLTRFOAERLARSHTQVNDPDAOSALITHGGMATLDAI
NRTPLALPLAFDQGVASRTIVSHIGIKRASRPTTSALARRRALITNDYQKRT
KIQTALRAGGMAADIVQAMRTQPVLAGSVATL"
3123..4271
/gene="crtY"
3123..4271
/gene="crtY"
/codon_start=1
/transl_table=1
/product="dycopene cyclase"
/protein_id="AAA21262.1"
/db_xref="GI:148396"
/translation="MRPYDILVAGGLANGLIARLQOQDPDKRILLIEAAPQAGN
HTWSPHADLTESQHRWAPLVVYHWYOVFPFTRKRLNSGIFSVTSQFAELQOR
KFGQHLWISRAVAEVAHADVRLNNGQVLSASAVIDRGYTNLSALNVGFQIQEMR
LSKPHGLSSPIMDATVDQNGYRFVSLPSATELLIEDTHYIDNATLEPERARONI
RDYAAQDQWQLOTLTREGALPILVTGDSATFMOQPLACSGLRAGLFHPTGYSLP

gene
LVALADRLSALDVFSSSIHQATIFFAHRMCOQRFRRMLRMLFLAGPADSRWRVM
ORFYGLPEDLISRFVAGKLTLDRLRLISGKRPVPVLAALQAIMTTHR"
4283..5761
/gene="crtI"
4283..5761
/gene="crtI"
/codon_start=1
/transl_table=1
/product="phytoene dehydrogenase"
/protein_id="AAA21263.1"
/db_xref="GI:148397"
/translation="MKPTVIGAGFGALAIRLQAAIPVLLIEGRDKPGRAYYE
DQGFDPADPVTITDPSAIEELFTLAGKQKDYVELPVAPEYRLCWSGKXFNVDN
QAOLEAOIOQPNRPDVEGYRQFLDYRAVAKGYLGTVPFLISFRDMLRAAQLAKL
QAMRTYVSQVASYIENDEHLROAFSPHSLVGNPFPATSSITLTHLEREWGWPFRG
GTRALYKGMITKIFQDGGGVVNAKYSHMETTQITAEVHLADGRFPTRAVANADV
VHTYRLDSQHPAAVQSKKLQTKRNSLPLVYFELNHHQDLNHTYFCGRYREL
IHFIENHDLADPFLYLHPAFLVDSLSLEEGGSYYVALPVPHLGTALDMVTGSPR
LRDRIPFYLEOHYMPGLRSOLVTORMTFPFPDQNLAAQGSFVSPEVLTOSAWERP
HNRDKTINMLYLVGAGTHPGAGIPVIGSAKATAGIMLEDLI"
5797..6687
/gene="crtB"
5797..6687
/gene="crtB"
/codon_start=1
/transl_table=1
/product="prephytoene pyrophosphate synthase"
/protein_id="AAA21264.1"
/db_xref="GI:148398"
/translation="MAVGSKSFATSKLPDAKTRSVMLYSKCHRCDDVIDQDTG
QTDQPSIQTPBEQRIMQLEMKTRQAVASQWHEPFAFAFEVANAHDIAVAVADHLEG
FAMDVREAEYIQLDPTLRVCYHAGVGVGMAQIMQVNDATLDRACDGLAFQLTNI
ARDIVDAHYGRCYLPASMLENGLKENYAAQENQALSRILRRLOBAEPYLCAT
AGLAGLPLSAAVIAATAKQYRKIGVKGVAQTEAMDHQSITTPPKLSILLMASSQA
ITSRMHPHRRPHLMORPI"
ORIGIN
Query Match 28.7%; Score 2486.2; DB 1; Length 6965;
Best Local Similarity 64.4%; Pred. No. 0;
Matches 3852; Conservative 0; Mismatches 2103; Indels 24; Gaps 8;
QY 1284 CTTGAACATTATTTGCTGCGCGGAGCAGAGCCGATGCGTTGGCGCGGCATGCGTCC 1343
DB 1004 CTTGAGCAGTTATGCTGTCGAGGAGAAACGGGATTTTGAAGTGGCGGAGATCGGAG 1063
QY 1344 GGACGCTGGACACCGGCAACGATTCGCGGCTGTCGCGCGGAGCAGCGGAT 1403
DB 1064 GGCAGCTGGACACCGGATTAACGATTCGCGGATGCTGTTGCTCAGACGACGCGAT 1123
QY 1404 ATGGGCTGTACGTCGCGCAGCAGAGCGGATCTTGATCTTGCTGTCGCGTGGAAATGGTG 1463
DB 1124 CTGGGCTGCGCTGTCACCGCAGGAGGATGCTGGATATGCGCTGCGCGCGTGGAAATGGTG 1183
QY 1464 CAGGCTGCTCATGATCTGTCGACGACATTCATCAATGATTAACGCGCGGATGCCACGT 1523
DB 1184 CACGCGGATCATGTTGATCTTGATGATATCCCTGTATGATATGAGGACGATGTCGGC 1243
QY 1524 GGGCGCGCGGCAATCCATCGATGAATATAGGGAACAGTGGCGATCTGCGAGCGGTGGCG 1583
DB 1244 GGCAGCTTACCGTTCATCGTACGATACGAGAGCATGTGGCAATATCTGGGCGGTGGCC 1303
QY 1584 CTACTCAGCGCGGCTTTGAGAGTGATTTGCCCTGCGCGCGGATCTGCGAGCAACGACAAA 1643
DB 1304 TTACTGAGCAAAAGCTTTGGGGTATTCGATGACATGAGCCGACGCGGCTGGCAAAA 1363
QY 1644 GCCAGACCATTTGCGAGCTCTCTCGCGGTGGGCTGCAAGGACCTGGTTCAGGGTCAAG 1703
DB 1364 AACCGCCAGTGTGAGATTAATCAACGCCATTGGTATGCAAGTTTGTATACAGGGTCAAG 1423
QY 1704 TTCAGGATCTGATGACGCGGACACAGCGGACGTCGGAAGGCATCACCTGACCAAT 1763
DB 1424 TTTAAGATCTGTCAAGAGGAGATTAACGCGCGCTGAGGCATTTTGTATGATACCAAC 1483

QY 1764 GAACGTAAAAACGAGCGCTGTTTCGGCCACAGCTGAGATGGCGGCGATTTGGGCGCAT 1823
DB 1484 CACTTTAAAAACGACACCTGTTTTCGGCTATATGAGATGGCTTAATTTGAGCGAAT 1543
QY 1824 GCGTCAGTGAAGGATGCTCAGCGCTTTAAGCTATTTTGGCGAGATTTAGGCTCAGGCTTTC 1883
DB 1544 GCGTCACGGAAGAGCGCGATTTACCTACACCGTTTTTCGCTGATATCTCGGCGCAAGCGTTT 1603
QY 1884 CAGTTACTGAGACGACTGGCGGATGGCTCTAAGACACACCGGCAAGAGCTGTCAACAGAT 1943
DB 1604 CAACGTGTGACCACTTAACGAGATGGCATGGCTGACACCGGTTAAGGACATCAATCAAGAT 1663
QY 1944 CAGGGCAAAATCCACGCGTGTGAGATGCTGGGCGCGGAGAGGGGTGAGCGGTGCTGCGC 2003
DB 1664 GAAGGCAAAATCGACCTGTGTCAATTTATTAAGCTTCAGAGCAAGTTGAAGACCGTCTGCGG 1723
QY 2004 GACCATTTAAGACAGCGCGGATGACACACTTTCGCTGCGCTGCCATCGCGGTGTGCGCAC 2063
DB 1724 CACAGTGTGATCTCGCGCAGTGAAGCACTCTCAGCGCGCGGTGTCAAAACGGTCAATTCAT 1783
QY 2064 CGTCAATATATGACGCGCTGTTTATCAACAGCTGGCGATGTTCAACTGAAGCGGTCA 2123
DB 1784 CAACAAATTTATCAGGCGCTGGTTTGAACAAAACCTCGCTCGCGTCAAGTAAAGATGCT-- 1841
QY 2124 TACCTAATGGGGCATTTTGGCGTTATTTGGCGCACCGCTCTAAGCACTTTCAAGCATTTGC 2183
DB 1842 ---GCAATGAGCCACTTTCGCGCGACATGCTCGCTTTTATACAGCAATGCTCGTGCATTCG 1898
QY 2184 AGGCGCTGGCGCAAAACGCTGTGGCGCGGAGACATTCGATCACTTTATTCAGCAAGTG 2243
DB 1899 AAAACCTCGCTCAGAGATTTAGTGTCTGCGGTATCAGTGACCTTTATTCAGCAAG 1958
QY 2244 ATGACGCGACCTTGTGAGCGGACGAGCGCATTTGCTTTGCGCGTGGCGGACGACGCG 2303
DB 1959 ACATTAACACCATTCATTCAGCATCCCAACATCGGATTTCAATGCTGAGAGAAAGAACGCC 2018
QY 2304 ATCTGCGCGATGCTCTCCAGCGAATCAAGCGGCTGGCGCGACCGGGGGGCTGTGCG 2363
DB 2019 ATCCGTGTGCTCATTTATTCGCGCTGTGCACTTAAGCGGCCCATCTCTGGGCGCGTCA 2078
QY 2364 TGTTCGCGTATTCAGCATCTGGCCAGACACACGATATGCTATGCGGGAATGTCGCG 2423
DB 2079 TCGTGGGCTCATATGAATGATGCGCGCACGACGATATGCTGTGCTGGCGTCTCCCA 2138
QY 2424 CCGTGTGACACGCGCTGACAGTGCATGAGCGTATTTGCCATCAATGAAGCGGCTGATG 2483
DB 2139 GAGCGTTTAACGATTTGGCGATGATGCTGTCAATGTTGATCAAGATGAGACGACGAGCG 2198
QY 2484 GTCTGTGTGAGAGAGGCTTACAGTGCCTTTCGTGTGCTGTGCTGTGCGCTGTGCGGTCA 2543
DB 2199 CTCTGTGTGCTGAAGCGCTGGGCTGCGGTTTGTCTGTGTGCTGTGCACTCCCGCTTA 2258
QY 2544 ATGCGGAGGCGGCGCATTCGCTGTGTGATGAGCGCTTTCGCTTTCAGAGATGAGAAG 2603
DB 2259 ATGCGGAGCGGATTTCCCTGCGGTGATGCTTTTGAATATGACACAGCGATGCGC 2318
QY 2604 CCGTGCACGCTATCAGCGACGAGTGAATCTAAGACCGCATGATGCTGTATGATGCG 2663
DB 2319 CAGGTAGCGCTATGACGACGAGAAAAATTTATGACTGGCTGATGCTGTGCGATGACC 2378
QY 2664 CTGTCAATGCTGTGATGCGCGCGCTTTCGCGCTGCGGAAAGCGCATGCTTATCATGAT 2723
DB 2379 GCGGTATTTGCCATCAACAGTCAACAGATGGGCTTGAAGACCCCGGAGTAACTCATCATG 2438
QY 2724 GTCTGTGCGCGCTGGGCAATCATGATGAGTGTGCGCGCTTTGATTTTCAAGCGACG 2783
DB 2439 GTTTTGTGCAATTTGCAAAATCATGATGATTTTATTCAGAACTGATTTTCCCGTAAAG 2498
QY 2784 AACTGCAGGCTGTATCAACGCTGGATCGCTGC--GGAATCAGATTGTAGCGCGC 2841
DB 2499 CGGTACGCGGCTTGTATGCTGTGTGGCGCGCTGTGCGAAGAACGACGACATCAACCT 2558
QY 2842 GCTGCGCGACCTGTGCGACGCGCTGC--CGACGCGGTGTATGCTCGCTGGGCGACGC 2900

DB 2559 CTTGCGCGCTTATTTTTCGCGCTCTGAAGAGCCCGGAATTTTCGCTGCTGGGACGCT 2618
QY 2901 TACAGGGGACGCTTTCGCTGTGCTGTGCTATGCTGCTAGGCTTCCGCAATCAGACG 2960
DB 2619 TGCAGGGTATCGCTACGGCTGTGTTAAAGCCATCTGTAAGAGCTGTGAAGATGATG 2678
QY 2961 TGTGCTGTGTGAGACACTGTGGGGGTTGACCGGCGACGCGACATCACTCAGAC 3020
DB 2679 GTACGCTTGTGTGAGCACTGTGGCGCTGTGCTCTTTCAGGCGGAAGACTGGCGC 2738
QY 3021 TGGCGGCTGTGCTGTGAGTACGATTTTGTGATCAGCGCGGCGCTGACAGATGCGC 3080
DB 2739 GAAGCGGTACACAGCAGTGTGTGATTTTTCGATCAGTGTGCGCGGCTGTGCAAGCGC 2798
QY 3081 AACTGTTTACTCAACGCGGTCTGAACAGTGGCTGGAAGCATGAGATGTGACGCG 3140
DB 2799 AGCTTGCATTTACCTACATGGCGGATGAATCGGTGTGAGTGCATTAATATTCAGCGC 2858
QY 3141 CGATGCTGGCGCTGCGATGCTTTCGATTCAGCGCGCGGCGGACGATTTGAGTGGC 3200
DB 2859 CCGTTTTCGCTTCCGTGTGCAATTTGATCAGCGCGGCTGTGCTCACTGATTTGTTCTC 2918
QY 3201 ACGGCGTGGCGCGCGGCTCTCAGCTTTCAGCGGCTGCGGAGCTGAGACACACTGCG 3260
DB 2919 ATGAGATGAGAAACGTGCTCTCCGTTTACACACAGCATGCGCTGTGCGCGCACATGC 2978
QY 3261 AACAGTGTGATGATGATGCTATGCTGTGCGATGCTGACATTCAGGCGAGCTGCG 3320
DB 2979 GTGCAATTTGTCACCAAGTGTGATACCGCGCAACGATGAGAAATTCAGACCGCTTC 3038
QY 3321 AGCGGCGGCTGTGATGCGCGCGGCTGATTTTGTGAGAGAGCGGCTGTTCAGACG 3380
DB 3039 GTTTAGCGCGGCGCAATGCGCTGACATGATGTTGTGAGAGCGGCTGTGACCGCGC 3098
QY 3381 AAATGCTGTGCGGAGGCGCACCTGATGCGCGACCTTATGATGATTCATTTGCTGCGTGC 3440
DB 3099 AGCCGTTTGAAGAGGAGCGGCT--ATGCGACCGCATTTATGATTTTATGAGGCGC 3157
QY 3441 CCGCTGTGCTAACGCGGCTGATTTGCGCTGTGCTTTACGCGACGCTGACCGCACTTAAGT 3500
DB 3158 CGGCGGTGCTAATGCGCTCATTTGCGCGCTTTCAGCAGCAGCACTGACATGCGTAT 3217
QY 3501 TTTGCTAATGAGATGACGCGGCGGCGGCGCAATCACTGCTGTTCATGCGCA 3560
DB 3218 TGTGCTTATGAGAGCGGCGCGCGCGGCGGTAATCACTGCTCTTTCAGACGCG 3277
QY 3561 AGAGTCAAGAGAGCGAGTGTGCTGTGCTGAGCGCGCTGCTTTCGCGCGCTGCGCGC 3620
DB 3278 TGAATTAACCGAAGCCACATGCTGTGAGTACAGCGCTGTGTATCATCTGCGCGA 3337
QY 3621 TTATCAAGTACGCTTCCCAACCTGTGCTGCGCACTGTGATGATTTGCTGATTC 3680
DB 3338 CTATCAAGTTCGCTTTCACGCGGCGTGTGAAGTGAACAGCGGCTATTTTCTGTGAC 3397
QY 3681 CTGAGAGATTTTTCGCGGCTTACAGCAGTGTGCTGTGCGCGCTTACGACCGGACG 3740
DB 3398 TTCGACGCTTTTTCGAGGCTTTTACAAACAGAAATTTGTGTGACACTTAATGATTTAGCCG 3457
QY 3741 GCGGCTCAGGAGTGTCAACCAACCGGCGGTCACTGTGCGGATGAGCGGATTTTACAGCG 3800
DB 3458 CCGGTGTGAGAGATTTCAACCGACGCTGTGATGATTAATGATGAGTGTGATGCTG 3517
QY 3801 GCAGGCGGTATTTGACGAGCGCGGCTGACCGGACACCGCATGTGCACTGCGTATCA 3860
DB 3518 CAGCGCGGTATTTGATGAGGAGGATTAACGCCAAATTCAGCGCTGAGCGTGGATTTCCA 3577
QY 3861 GGCATTTTGTGCTCAGAGTGTCACTGTGCGCGCGGCTGCGCTGACGACGCAATATT 3920
DB 3578 GCGCTTATGACAGAGATGAGGAGCTCACTTAACCGACGCTTATCTTACCCATTA 3637
QY 3921 GATGAGCGGAGGTGATGACGAGCGGTTATGCTTTTATCAACCTGCGCGCTCAG 3980

D 3538 TATGATGCAACGCTGCATGCAAAATGTTATCGCTTTGTTTAAAGCTCCGCTATC 3697
Q 3981 TGCCAGCGTTTACTGATTTGAAGATACCCACTAGATCAACCATGCAACGCTGAATGCGC 4040
D 3698 TGCACAGAACTGTTAATCGAAGATACCCACTATATGATTAAGCGACACTGAAACCTGA 3757
Q 4041 ACAGCGCGCGCTACATTAAGATTAAGCCACAGCGCGCGCTGGAATTTGGCCAGCT 4100
D 3758 ACGCGCGGACAAATATTCGCGATTAATGCTGCCAGAGGATGGAGCTTCAGACTCT 3817
Q 4101 GCTCGCGAGAGAGACGCGCTGCTGCGATCAGCTCAGCGCGATATCATGATCTG 4160
D 3818 GCTTCGTAGAGGACAGCGCGCATTTGCGAATACCTTAACGCGCATAGCAACGGGTTTG 3877
Q 4161 GCAAACAGACAGCGGACACCTGTGACGCGGCTGCGCGCGGACTGTTTACGCCACAC 4220
D 3878 GCACACACAGC---CTTTAGCCTGACAGTGCCTGCGCGCGGATGTTTCAACCCACAC 3934
Q 4221 CGGTTACTGCTGCCCGCGCGGTGGCGCTGCGGAGAAATGCGACAGCAGCTGCCCGC 4280
D 3935 CGGTTATTCACCTCCCTTGGCGTTGGCTGCGATGCGCTGAGTGGCTTGAATGCTT 3994
Q 4281 CGAGCTCACACGCTGAGCCACTGCATCGAATCCTTTGCCGTGACGACTGCGCGACA 4340
D 3995 TACATCGTCTCAATTCATGAGGCGATTACCCACTTTGCCACGAGGCGCTGGACAGCA 4054
Q 4341 GCGCTTTTCCGCTGTTAAATCGATGCTGTTTCTTGGCGGACGCGCTGAACAGCGCTG 4400
D 4055 ACGCTTTTCCGCAATGTTGAATCGCATGCTGTTTTCGCGGACCGCGCGATTCACGCTG 4114
Q 4401 GCGGTAATGACGAGGTTTTCACGCGCTGACGCGCGATGATTAAGCGCTTTTACGCGCG 4460
D 4115 GCGGTTATGACGCGCTTCTATGCTTACCTGAAGATTTAATCTCCGTTTATGCGGG 4174
Q 4461 GCAACTGCGCTTCAGCGATTAAGACGCAATCTGTGCGGCAACCGCGCGCTCTCG 4520
D 4175 CAATTCACGCTGACCGATCGGCTACGTAATCTGAGCGGCAACCGCGCTGTTCCGTAAT 4234
Q 4521 CGAAGCGCTGCGCATTTGATGATGACCTCTCCGTTACAGGAGAAATTAATGAACGC 4580
D 4235 AGCGGCAATGACGCAATTAAGACATCATGTTAAAG---GCAACATATAACCA 4291
Q 4581 ACCTATGATTTGCGCAGGCTTCGCTGCGCTGCGCTGCGATTCGTTGCAAGCGGC 4640
D 4292 ACTACGGAATTTGATGAGGCTTTGCTGCGCTGCGCTGCGATTTGCAATTCGTTGCAAGCGGC 4351
Q 4641 GCGGTCGCGTACGCTGCTGGAACAGCGCGATTAAGCTGCGGCGCGCTATGCTAT 4700
D 4352 GGGATCCCTGCTTACTGCTTGAACCAACGCGAACAACCGGTCGCTTATGCTAT 4411
Q 4701 CAGGATCAGGCTTTTACTGTTGATGCGCGTCCGAGCGGATTAACGATCCAGCGCTATC 4760
D 4412 GAATATCAGGATTCCTTTGATGACGAGGCTTACCGTTATTCAGATCCAGCGCAATT 4471
Q 4761 GAGCGCTGTTTACGCTGCGCAGCAAGCACTCATGATTAATGCTGACCTGATCCGCTG 4820
D 4472 GAAGAACTCTTACCTGCGCGGAAACAGTTAAAGTTTACGTTGAACTGCTCCGCT 4531
Q 4821 ACGGCAATTTTATCGCTGCTGGAAGACGCGACGCTGACCTAGACACATATG 4880
D 4532 GCGCGCTTTTATCGCTGCTGGAAGACGCGAAGGTTTCAATTAAGTAAGATCAG 4591
Q 4881 GCGGAGCTGAGACAGATTGCACTTTTAATCCCGAGATGCGCGGTTTACCGCGAG 4940
D 4592 GCAACGCTTGAAGGCGACATTCAGCAATTTATTCACGCGATGTTAAAGCTTATGCTAG 4651
Q 4941 TTTCTGCGCTAATTCACAGATGTTTCTGTAAGGCTATGAAACTGCGGACCGTACT 5000
D 4652 TTTCTGCGCTAATTCACGCGGCTGTTAAAGAGCTATCGAAGCTTGGACGCTGCT 4711
Q 5001 TTTCTGCAATTTCCGCGACATGCTGCGCGGCAACGCTGCGCTGCGAGCGCTG 5060
D 4712 TTTCTGCTGTTCAAGGATATGCTTGGCGCGCGCCCACTGCGAAACTGCGAGCATG 4771

Q 5061 CGCATGTTCAACAGATGTTGGGAAATTTATCATACATATCTGCGCAGGCTTT 5120
D 4772 CGAACCTTTTACAGTAAGTTTGGCAGCTTACATTAAGGATATAATCTTCTGCGGCTTT 4831
Q 5121 TCCCTTCACTCGTTGCTGAGTCCGCGGTATCTTTTTCAGAGCTTTCGATCTTACTTA 5180
D 4832 TCTTTTCACTCGCTGTTAGTGGGCGCAATCCCTTGGCCACTCTCTCATTTTATGCTA 4891
Q 5181 ATTACGACATGAGACCGCAATGGGCGTGTGTTTCCGCGCGGATACCGGTGCGCTG 5240
D 4892 ATACATGACATCGAAGCGAATGGGCGTCTGTTTTCACGCGGTGATACCGGGCATTA 4951
Q 5241 GTTGATGAGATGCGCGGCTGTTTCCGATTTTGGGCGGTAACTGCTGCTCAACGCCGA 5300
D 4952 GTGAAAGGATGATTAAGCTGTTTCAAGATCTGGGTGGCAAGTGTACTGAATGCGAG 5011
Q 5301 GTCAGCAGCTGAGACCGAGGTTAACCGCATCAGCGGTTCAGCTGAAGATGCGCGC 5360
D 5012 GTACGCGCATGGAACACAGGGGATACCATTAAGCCGTGATTTAGAGACGACGCG 5071
Q 5361 CGTTTTCGCGCGCGCGCTGCTGCTCAATGCTGACGCTGTGATACATCGATGCTG 5420
D 5072 AGATTCGCAACCGGCGCTGCTGCTCAATGCGAATGTGTTCAACCTATCGGACCTG 5131
Q 5421 TTTAAGCCAGATCTGCGCGCGTAAACGCGCGCAACGCTGAAGCCGAGATAGC 5480
D 5132 TTTAAGTACGATCTGCGCGCTGAAGCAGTCAAAAACCTGAAACTAAGCGCATAGC 5191
Q 5481 AACTCGCTGTTTACTTATTTTGTGCTTAATCATATGCCACCGCAGCTGCGGACAC 5540
D 5192 AACTCATGTTTGTCTTATTTTGTGCTGAATCATACATACGATACGCTTGGCCACAT 5251
Q 5541 ACGGTGCTTTGGTCCGCGCTATCGTAATGATGATGATGATCTTCAATAGACGAG 5600
D 5252 ACGGTTGTTGCGCGCTCCCTACCTGACGCTATCATATAATTTTCAATCATAGCGA 5311
Q 5601 CTGCGGAAATTTCTGCGTATCTGATGATGCGCGCGCTGCAAGCATCCGCTGCGGCA 5660
D 5312 CTGCGAACGATTTCTACTTACTGACGACACTTGTGTACAGATTTGTGACTGCGCA 5371
Q 5661 CCGCGCGCTGCGGCAAGTTTTCATGCTGCTGCGCGCGGTCGCGATCTCGGTACCGCGCA 5720
D 5372 CCGGAAGGTTCCGCGAGTTACTATGTTAGTCTCCGTCACATTTAAGCACAGCGAAC 5431
Q 5721 ATTGACTGCAACAGAAAGGCGCGCTTCCGCGATGCAATTTTGTCTTATCTGAGAG 5780
D 5432 CTGACTGACGCTGTAAGGCGCGCACTGCGGATCGTATTTTGAAGTACTTGAAGCAA 5491
Q 5781 CACTATATGCGCGGCTGCGACAGAGATTAGTACACACGCTATGTTTACGCGTTGAT 5840
D 5492 CATTAATGCTGCTGCTTGGCTGACCACTGTTGATTCGAGGAATGTTTACCGCTTCAT 5551
Q 5841 TTTTCGACACGCTGCAACGCGCATCAAGGCTCAAGCTTTTTCGTCGAACCATTTTGAAG 5900
D 5552 TTTTCGACACGCTCAATGCTATCAAGGCTCAAGCTTTTTCGTTGAGCCGTTCTACA 5611
Q 5901 CAAGCGCTGTTTCCGCGCGCATTAACCGGATGCGGACATTAACCTTTATCTGCTG 5960
D 5612 CAGAGCGCTGTTTCCGACGCAACCGGATTAACCTTTATCTTCTGCTGTA 5671
Q 5961 GGGGCTGCGACGATCCCGGTGCGGCTGCGAGCGTGTATCGGCTCCGGAAGCGAC 6020
D 5672 GTGTGAGGACCATCCCGCGCGGCTATCCGCGCTGATTCGCTGCGTAAAGCACT 5731
Q 6021 GCCCACTGATGATGAGATCTGACCGATGAACCAACCGCGCTGATTAAGCAGGTC 6080
D 5732 GCAAGTTTATGCTGAGGACCTGA---TTTGAATTAATCGTGTACTCAATCATAGCGG 5788
Q 6081 CGCAACCATGCGCGAGGCTCCAAAAGTTTCCGACGCTTACCGGCTATTTGATCTT 6140
D 5789 TCGAAACGATGCTGTTGGCTTCGAAAAGTTTTCGACTGCTCAAAAGTTTGAACGAA 5848

QY	6141	CAAGCGCGCGAGTACGCTGATGCGTGTACGCGTGGTGTGCTGACCTGTGAGCATGTGATG	6200
Db	5849	AAACCCGGCGAGTGTACTTATGCTCTACTCTCTGGTGCAGCATGTGACCACTGATG	5909
QY	6201	ATGCTCAGACCCCTGGCGGAGGCGGCACGCGACGCGCGGTGGCGGATGTCACAGGCGCGGA	6260
Db	5909	ACGATTCAAACCTTGGGTTTTCAAACCGATCAGCCGCTACTTTCAAACCGCTGAGCAACGCC	5968
QY	6261	TGCGCCACTTGCATAATGCAAAACCGCGCGCGCTTACAGCGGTGCCCACTGATTAACCG	6320
Db	5969	TGATGACAGTTGGAATGAAAACGCGCGCGCGCTTATGCGCGATTCGCAGATGACGAACCGG	6028
QY	6321	CGTTTCGTCCTTTTCAAGGAAGTGGCGCTGACGCAATCAGCTTCCCAGACGTGGCTTTTG	6380
Db	6029	CGTTTGCTGCGCTTTTCAGGAAGTGGGGAATGGCCCATGATATTTGGCGCGCTTACGCTTTG	6088
QY	6381	ATCATCTTGAAGGGTTTSCGATGATGCGCGCTGAGAAAGTTATGTCGTTTCCGGGACGA	6440
Db	6089	ATCATCTTGAAGGCTTTCGATTCAGCTTGATGTGCTGAGCGGAGTACATTTCAGTTGAGCAT	6148
QY	6441	CGCTGCGTTACTGCTATCAGTGGCGCGCGCTGTGGGGTTAATGATGGCGCGCGTATGG	6500
Db	6149	CATTACGGTATTTGCTATCAGCTTGGAGCGGTGTGCGCTTATGATGATGGCGCAGATTTATGG	6208
QY	6501	GCGTACGCGATGACGCGCTACTTCGATCAGCGCTGTGATTTTGGGTCTGGCGTTTACGTTA	6560
Db	6209	GCGTACGCGATGACGCTTACCTCTGACCGCGCTGTGACTTTGGCGCTGGCATTTTCACTTGA	6268
QY	6561	CCAAATATCGCAGGGATATGCTGTTGAGAGCGCGGAGAAAGGCGCTGTGCTATCTGCACAAA	6620
Db	6269	CCAACTATGCGCGCGGATTTGTGGAAGATGCGCATGTGGGTCTGTTTATCTGGCGGCGGA	6328
QY	6621	GCTGCGTGGATGAGGCGCGGACTGACGCGCGCCACGCTTGCAGATCCGCAATCGCGCAG	6680
Db	6329	GTTGGCTGGAATAATGAAGCATGAATAAAGAAATTTATGCGGCACCGGAAACCGCGCAGG	6388
QY	6681	CGCTGCGCGCGCGCTGGCAGCGGTCTGTGTGCGGAGGCGGACCGTATCTATCAGTCAAGCG	6740
Db	6389	CGCTGAGCGCGCATGCGCCGAGCGTCTGTGTGACAGGAAGCCGAACCTTATCTATTTGTGCGCTA	6448
QY	6741	GCAGCGGGCTGCCAGGATTTGCGCGCTCGGTTGCGGCGTGGCGCATGGCCACCGCGCGCGCG	6800
Db	6449	CGGCGGGACTGGCGCGGCTGCGCTTGGCGCTCAGCGCTGGCCCATGCGCACAGCGAAGCAGG	6508
QY	6801	TTTACCGGGAATTTGCGCTAAAGTGCAGCATGCGCGTGGCGCGGCGCATGGGATACGCGCC	6860
Db	6509	TTTATCTGTAATTTGGGGTGAAGGTGAAACAGCAGCAAGAACAGAGCTCGGGAACCATGCGC	6568
QY	6861	AGCGCACCATGTAAGGCGAAGAACTGCGCGCTGTGTGTGAAGAGTGC CGGCTCTCGCGCTTA	6920
Db	6569	AGTCAACACACACACCCGAAAAATTTATGCTTGTGATGGCATCCGCTCAGGCGCATTTA	6628
QY	6921	CTTGCGCGCTGTGCATCCCGAGGCGGCTCTGCGCGGTCTGTGAGAGGTGCCGCTTAC	6980
Db	6629	CTTCCCGGATCGGCTCATTCGCGCTGCGCTGCGCTGCATCTCTGGAGGCGCCGAAATTAGC	6688
QY	6981	ACGACGCGCATGCGCTGGCGCAGCGGTGCGCTGACGCTTGTGTCAACGCTGGGCGGTAAAG	7040
Db	6689	-----CGAATGTGCTTCTCTTAAAGTGTGCGCTGAAAGTTTATATAGCGCGCGCGCATAGAG	6742
QY	7041	AAAGCCAAAGAAAGCAGAGCTTCCCGTCCCGGACCGCATGATGATCGGTGGCGCAT	7100
Db	6743	AAAACCAAAAGAAACACAGCGCTTCTTTCGCTTTTACCGCATGATGATGATGCGGTGGCGCAT	6802
QY	7101	ATPACACCGCTTCAGATGACCTTTGCGGGAATATAGGAAACGGCGCAGCGTTATGTGAC	7160
Db	6803	ATPACACCGCTTCAGATGACCTTTGCGGGAATATCCGGCAACGGCGCAACCTGATGTGAC	6862
QY	7161	CAGGCGCTCATGACCATATAATAGAGAGCCCATAGGTCTGTCAATGCGAGGCGCAATTCGA	7220
Db	6863	CAGCGCGCTGTGAACCATGAAATAGAGAAATCATAGAGCGCTCATGCGCGCACCAATTCGA	6922
QY	7221	CTGACGCGCGCAGACGCGCTGTGACACCGCAATTCAG	7259

```

Db          6923 CTGAAGCGGCACATTCCTCCGTAACGCCCAACAATAAATAG 6961
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
RESULT 6
ERMCRS
LOCUS      12753 bp      DNA              linear    BCT 11-APR-2001
DEFINITION
Panacea agglomerans pyrophosphate synthase (crte), zeaxanthin
glucosyl transferase (crtx), and lycopene cyclase (crtl) genes,
complete cds; phytoene dehydrogenase (crtl gene, partial cds;
phytoene synthase (crtb) and beta-carotene hydroxylase (crtz)
genes, complete cds; and unknown genes.
VERSION     M87280.1 GI:148404
KEYWORDS
ORGANISM   Panacea agglomerans
            Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
            Enterobacteriaceae; Pantoea.
REFERENCE   1 (bases 1 to 12753)
AUTHORS    Hundle,B., Alberti,M., Niveststein,V., Beyer,P., Kleinj,H.,
            Armstrong,G.A., Burke,D.H. and Hearst,J.E.
TITLE       Functional assignment of Brwnia herbicola Eho10 carotenoid genes
            expressed in Escherichia coli
MOL. GEN. GENET. 245 (4), 406-416 (1994)
JOURNAL    JOURNAL
MEDLINE    PUBMED 95107236
            7808389
2 (bases 1 to 12753)
AUTHORS    Armstrong,G.A., Hundle,B.S. and Hearst,J.E.
TITLE       Evolutionary conservation and structural similarities of carotenoid
            biosynthesis gene products from photosynthetic and
            nonphotosynthetic organisms
Meth. Enzymol. 214, 297-311 (1993)
JOURNAL    JOURNAL
MEDLINE    PUBMED 93225848
            8469144
3 (bases 1 to 12753)
AUTHORS    Hundle,B.S., O'Brien,D.A., Beyer,P., Kleinj,H. and Hearst,J.E.
TITLE       In vitro expression and activity of lycopene cyclase and
            beta-carotene hydroxylase from Brwnia herbicola
FEBS Lett. 315 (3), 329-334 (1993)
JOURNAL    JOURNAL
MEDLINE    PUBMED 93138098
            8422926
4 (bases 1 to 12753)
AUTHORS    Armstrong,G.A., Alberti,M. and Hearst,J.E.
TITLE       Conserved enzymes mediate the early reactions of carotenoid
            biosynthesis in nonphotosynthetic and photosynthetic prokaryotes
Proc. Natl. Acad. Sci. U.S.A. 87 (24), 9975-9979 (1990)
JOURNAL    JOURNAL
MEDLINE    PUBMED 91086634
            2263648
5 (sites)
AUTHORS    Hundle,B.S., O'Brien,D.A., Alberti,M., Beyer,P. and Hearst,J.E.
TITLE       Functional expression of zeaxanthin glucosyltransferase from
            Erwinia herbicola and a proposed uridine diphosphate binding site
Proc. Natl. Acad. Sci. U.S.A. 89 (19), 9321-9325 (1992)
JOURNAL    JOURNAL
MEDLINE    PUBMED 93028456
            1409639
FEATURES
source     Location/Qualifiers
            1..12753
                /organism="Panacea agglomerans"
                /mol_type="genomic DNA"
                /strain="Eho10"
                /db_xref="taxon:549"
                /ligase_1fb="ATCC 39368"
                complement (<1..1588)
                /note="ORF2"
                /citation=[1]
                /codon_start=1
                /transl_table=1
                /product="unknown"
                /protein_id="AA64974.1"
                /db_xref="GI:551854"
                /translation="MSFLIHNNHNIEMRFROILRLPEKMKSLRYVATCAPFTSL
HALTPSCAEIIRDAETITLIRNGNSAVLNELGQALPTNORELLLEEMRWRSIV

```

GCSAEKQDRNSQASVPEFYQSSIGTGLIMNNPAAIGESAFAGKHSQMSRTVQSDNYIT
 KCSGIELAKDRNSQASVPEFYQSSIGTGLIMNNPAAIGESAFAGKHSQMSRTVQSDNYIT
 TAGSDVMEITRQYAKATGTPPAAPAEFFISGLQCKRKYETQOEDVDAEYHRRLPLPS
 VNVVIDFEMHPQGNVLCFDPDVPMDPRAVWEELBSGILMWSWPEYEAARPLPEPOKAK
 AGMGTIVSRGYNVLCDEMGNTFFPDATHPQAPQFVPMPTVAKNFYDYGIKLFMYDEABA
 PEYRAVDPEFNRYHAGPVLEVRNRYPRPFAQFYGLAEENEIVNLVRSNMGASOR
 YGLVMSGCVHSSPFSFNOLAAAGLNMGLAGIPWMTTIDGFGGSHIDPFAHELLIR
 MWMQVFCVCVLMHGYREBRJ"
 complement(1596..2414)
 /note="ORF3"
 /citation=[1]
 /codon_start=1
 /transl_table=1
 /product="unknown"
 /protein_id="AAA64975.1"
 /db_xref="gi:148406"
 /translation="MLTKNNLMIVQALRQHRMLCLFGVAYIFSSVGNGLTQTLILG
 QLLRAHAPSTLTLLTMLATLPGFISGLVIGMMFVPTYLGSASVFMMLLAHLIGKRA
 AGSLIAKRLTRNFVKVQIFGYCAVLGAVLSLIPAPKSPVLVPLFEIISTLQATPTT
 TLMYVMDADVADYEGMSOGKMDGIIFSTPLAVLKGALSGAIVGMPLIGSGVYANA
 POTSIAHNCIYALFTVYVGLLSLAAPATLRYKLDQDTQMEINLAKMGIIVKS"
 3140..3451
 /note="ORF4"
 /citation=[1]
 /codon_start=1
 /transl_table=1
 /product="unknown"
 /protein_id="AAA64976.1"
 /db_xref="gi:148407"
 /translation="MAIRSLGTTGHTHPVYFGNFGWTYDEQOSRSLDALLERGI
 NAITDADVYSAAPNKGSESETILGKFAHPGAREKITLFTYQVSLKARMLRTRM
 M"
 3535..4458
 /gene="ctxT"
 3535..4458
 /gene="ctxT"
 /citation=[4]
 /citation=[2]
 /codon_start=1
 /transl_table=1
 /product="pyrophosphate synthase"
 /protein_id="AAA64977.1"
 /db_xref="gi:148408"
 /translation="WVSGSKAGVSPHREIEMWROSIDHLAGLLPETSQDIYSLAR
 EGVMPGRIPLMLLAARDIRYQSGMPTLLIDACAVELHTASLMLDDPQMDNBE
 LRRQPTTHKKEGSVAIIASVGLSKAFGLIATGDPERRRAQVANEISLAVGQVQ
 IYLCGFRLNDPAALDRTPDALISTHMLTGILFSAMQIYVIAASPSSTETLHAAPA
 LDFGAFQPLDRLDHPETGDKRMKDGKSTLVNRLGADAARQKLRHIDSAKHILT
 PACPGAGIROMMHPGHLADMSPVWIKIA"
 4521..5564
 /note="gene not found in *Erwinia uredovora* crt gene
 cluster; ORP6"
 /citation=[1]
 /codon_start=1
 /transl_table=1
 /product="unknown"
 /protein_id="AAA64978.1"
 /db_xref="gi:148409"
 /translation="WKDSRLVQRKNHDILVLPBRAVYQASAGFERMRFTCHALPEL
 NFSDTLTETFLNRQLQAPLILSSMTGQVBSRSHNRHLABAQVLIAMGVSGQRAVA
 IESDGLGIDKTLRQLADPVLNLAAGCAOTGRKGIYARAVEMTEADLIYHNPML
 LOEALOPGQDNRMRGLAIETLVRELVPVLVREKVGASIRTVAGOLIDGVTGVIDV
 AGAGTSGVAAGGERAATEOORSVANVFADMGIPAEALNDIAEMPMPLIAGSGIKK
 NQVDAKALRLGACNVQDAAVLGSAGVSTERYIDHNVITIEQLRVACFTGSRSLSD
 LKQADIRYVRP"
 5561..6802
 /gene="ctxX"
 5561..6802
 /gene="ctxX"
 /citation=[3]
 /citation=[1]
 /codon_start=1
 /transl_table=1

[illegible]

Qy	2196	AAACGCTGCTGGGGCCGGGACATCGCATCAACCTTTATTCAGAAAGTATGACGACACT	2255
Db	5628	TGAGATGGGCCCAACGGGGCCACCGGGGTGACCTTTCTC-----ACGGCAACGTCGGCT	5681
Qy	2256	TGCTGAGCGACGAGCGCATTCCTTTGTGGCCGCTCGCGAGCGACGCACTCTCGCGGAT	2315
Db	5682	CGCTGGGAGAGACAGAAACCGGACCGGGTGGCCCTTATTCATCTTCCCGCAGCGTGCAAC	5741
Qy	2316	CGCTTCACAGCACTCAGGGCGCTGGCGCACCGGGCGGGCTGTGCTGTTTCGCTGA	2375
Db	5742	AGGCCCCAGCCCAACGTCACACAGAGTAAACGGCAACCTGCTGGCGGTGATTTGGCGCA	5801
Qy	2376	TTACAGATCTGGCCAGACACACGATATGCTATGCGCGCAACTGCCCGCGGGCTGGCAAC	2435
Db	5802	TGTCATCCCTG-----ACGATGTGCTTGCCAGAGATTGCCCGCTATTTCAACGC	5852
Qy	2436	GGCTGACAGTCGATGCGGTGATTTGCGCATTAATGGAAGCGGCTGGTGTCTGTGGCAG	2495
Db	5853	GGCTGGGGGTGGACCGGCTATTTGTGATGATGATGAGCCCGCGGAAGCTGTGTCCCG	5912
Qy	2496	AGGCGTTACAGCTGCGCTTGCTGTGGGTGGCGCTGGCGGCTGCAATGCGCAAGCGG	2555
Db	5913	AGGCGTGGGACATCACTTATCTTATTTGCTGCGCGCTCGCGGTCAACCGCAGCGCG	5972
Qy	2556	CCATTCGCGCTGGGTGATGATCCCTTTTCGCTTTCCTAGATGAGAAAGCGCTGGACGCT	2615
Db	5973	GTCTGCGGCTGGCGGTATGCGCTTTCATCAACGCCAGGATTAAGAAAGCCCTGGCGGCT	6032
Qy	2616	ATCAGGCGCAGAGTACATCTACGACCGCATCAATGCTGTGTCATGCGCTGCTC	2675
Db	6033	TTCAAGTTCAGCGAAGCATCTTACGATGCGCTATGATACCGACCGGGCAGAGATCTTGC	6092
Qy	2676	GTCAATGCGCGGCTTTCGCGCTGCGCGAAGCGCATGATGATTAATCAATGCTGTGCGCG	2735
Db	6093	GCACGCGCCAGCGCTTTGTTTGCGGAGGACGAGGCGTCTCGACGATGTCTCTCGCGCG	6152
Qy	2736	TGGGCGCAATCACTGACGCTGTGCGCGCTTTTATTTTCAAGCGCCAGCAACTGCCAGCT	2795
Db	6153	TGGGCGAGATTAGACATTCGTTTCGGCCCTCGACTTCCACGCCCGGGCGCTCGCGAACT	6212
Qy	2796	GCATTCACAGGTTGGGTGCGCTGGAGCTCAAGTTGCTAGCGGCGGCTCGCGCACCT	2855
Db	6213	GTTTTCACTACGTGGAGACCTGCGCTATC-----AGCCGCGCGCGCAGGTGAAC	6263
Qy	2856	GGCCAGGCTGCGCGACCGCGGTGATGCTGCTGCTGGGACAAGCTACAGGGGCAATCGT	2915
Db	6264	GCTGCGCACGAGAGCGCGGGATCTTTGCTGCTGCTGGGCAACCTCCAGGGGCAACGCTC	6323
Qy	2916	TTGGCTGTTTTCTGCACTGTGGCTCAGGCTGCGCGAATACAGACGCTGTGCTGTGGTGG	2975
Db	6324	TACGCTGTTTTAGAAAGATCGCGCGCTGTGCTCAGCGTGGGGCGGAGGTGACCATTTG	6383
Qy	2976	CACACTGTGGCGGTTTACCGCGCAGCGACATACGTCAGACTGGCGGGTGTGCTGCT	3035
Db	6384	CCCACTCGATGGCTTGACGCCGCCGCCAGCGCATCTGCTTACGCTTGCGCGCGGAGCGG	6443
Qy	3036	GGGTCGACGATTTTGTGATCAGCGGGCGGCGCTGACAGATGCGCAACTGTTTATCACTC	3095
Db	6444	AGGTGTCACGCTTTGTGACACACCGCGCTACGTTGCCAGAGGTATCTGTGTATCAACC	6503
Qy	3096	ACGCGCGGTCTGAACAGTGGCTGGAAGACTGGAATGTGGCACCGGATGTGGCGCTGC	3155
Db	6504	ACGCGCGGTCTCAATACGCTACTGGAATGCGCTGCTCCGCAAGCGCGGTGCTGGCGGTGC	6563
Qy	3156	CGATTCGCTTCGATACACCGGGGTGGCGGCGCATTTAGATGGACACGGGCGTGGCGCGG	3215
Db	6564	CACCTCTTTTGACACACCGCGGTGGCTGCGCGCGCTGCTATTAACGGGGCTGGGTGCGC	6623
Qy	3216	GCAGCTCAGTTTTACGCGGGGTGCGCGACCTGGAGACCACTGCAACAGTTTCTGATGTG	3275
Db	6624	GGGATATGGCGTTTGCAGACAGAGACGCTGGCGGATGAGATTTGCCCAACTGTGGGGG	6683
Qy	3276	ACGATTCGATCTGCTGCGATGTACGCAATTAGAGCGCAGCTGCAAGCGGGCGGGTGCCT	3335

Dh 7740 GCTGAACCGAGTCTTTTCTGCGCGGCGGAGAGAACCGCTGCGGGTATGACGCG 7799
Qy 4416 TTTTACCGGTTACGCGGATGATTAGCGGCTTTTACCGCGGCAACTGCGGCTTACG 4475
Dh 7800 CTTTATAGGGCTGCGGAGCCACCGTAGAGCGCTTTTACCGCGGCTCTCTCTT 7859
Qy 4476 CGATAAGCAGCATCTGTGCGGCAAAACCGCGGTCCTCTCGCGGCAAGCGTGCAGC 4535
Dh 7860 TGAATAAGCCCGCATTTTGAAGGGAACCGGCTTCCGTGGGCGAAGCGTGGCGGCG 7919
Qy 4536 ATTGATGATGACCTCTCGTTTACGAGGAGAAATATGAAAACGACCTATGATTTGCG 4595
Dh 7920 GGGCTGAACCATTTTCTGACAGACGATTAAGATGAAAAAACCGTTGATTTGCG 7979
Qy 4596 GCAGGCTTCGCTGCGGCTGCGGCTGCGGCTTGTGCAAGGCGCGGCGGCTGCGGCTACG 4655
Dh 7980 GCAGGCTTCGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTTCCAAACGTA 8039
Qy 4656 CTGCTGGAACAGCGCATGAGCTGCGGCGGCGGCTTATGATTCAGATCAGGCTTTT 4715
Dh 8040 CTGCTGAGAGCGCGGAGACAGCCGCGGCTGCGGCTTACGCTGCGCATGACAGGCTTT 8099
Qy 4716 ACCTTTATGCGGCTGCGGCGGATTAACCGATCCAGCGGCTATGAGCGGCTTTACG 4775
Dh 8100 ACCTTTAGCGCGGCGGCGGATGATCAACGATCTTACCGGCTTGAAGCGCTGTACCC 8159
Qy 4776 CTGCGAGGCAAGCATCAGTATGATTTATGCTGATGCGGCTGAGCGGCTTTTATGCG 4835
Dh 8160 CTGCGCGGAGCGCATGAGGATTAAGCTGCGGCTGCGGCTTACCGA 8219
Qy 4836 CTGCTGAGGAGAGCGGAGGAGCTGAGCTGAGCAACATCAGCGGCTGAGAG 4895
Dh 8220 CTCTGCTGAGAGTCCGGGAACCCCTGCACTATGCTAACAGACGCGGCTTGAAGCG 8279
Qy 4896 CAGATGCGCACTTTTAAATCCCGAGATGTCGCGGTTACCGGCACTTTTCTGCGCTATTA 4955
Dh 8280 CAGATTAACCGCATTAACCCCGGAGCGTACGAGGCTTACCGGCTTCTGCTTACCTCC 8339
Qy 4956 CAGATGCTTTTCTGAGGCTATCTGAATGAGGAGCGGCTTCTTCTGCTTCTGCGC 5015
Dh 8340 CAGGCGGATTCAGAGGAGATTTTGGCTGCGGAGCGGCTTCTCTCTTCTGCGC 8399
Qy 5016 GACATGCTGCGGCGGAGCAGAGCTGAGGCTGAGCGGCTGAGCGAGCTGATACAGC 5075
Dh 8400 GACATGCTGCGGCGGAGCAGAGCTGATGAGCTTACAGGCGGCTGAGCGGCTTACAG 8459
Qy 5076 ATGATGCGAAATTTATTCATGAGCATCTGCGGAGCTTTTCTTCTGCTGCTG 5135
Dh 8460 TCGGTTTCGCGCTTATTTAGAGATGAGCATCTGCGGAGGCTTCTGCTTCTGCTG 8519
Qy 5136 CTGCTGCGGCTTATCTTTTGAACGCTTTCATCTTATCTTATTTACAGGCTGAG 5195
Dh 8520 CTGATGAGCGGAGAACCCCTTACACCTGCTCATTAACACCTGATTCAGGCGCTTGA 8579
Qy 5196 CGCAATGAGGCGGCTGCTTCCGCGGCGGATCCGCTGCGGCTGCTGCTGCTGCTG 5255
Dh 8580 CGGAGTGGGCGGCTGCTTCCGCGGCGGAGCGGAGCGGCGGCTGCTGCTGCTGCTG 8639
Qy 5256 CGGCTGCTTTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 5315
Dh 8640 AAGCTGTTTACGATCTGCGGCGGAGATCGATCTCAACCGCGGCTGCAAGAGCTG 8699
Qy 5316 ACCGAGGCTAACCGCATCAGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 5375
Dh 8700 GTGCGCATTAACCGCATCAGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 8759
Qy 5376 GCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 5435
Dh 8760 GCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 8819
Qy 5436 GCGGCGGCTTAACCGCGGAGCGCTGAGGCGGAGCGGATGAGCAACTGCTGCTTGT 5495
Dh 8820 GTGGGCGAAGCGGCGGAGCGCTGAGGCGGAGCATGAGCAACTGCTGCTTGTG 8879

Qy 5496 CTCATTTTGGTCTTATCAATGCCAACCGGAGCTGCGGAGCAACGAGTGTGCTTGT 5555
Dh 8880 CTCCTACTGCGCTGAACCAAGCTTATTCAGCTGCGGAGCAACATCTGCTTGT 8939
Qy 5556 CCGGCTATCTGATTAATGATGATGATCTTCAATAGCAGCGGCTGCGGAGATTTT 5615
Dh 8940 CCGGCTATCTGATTAATGATGATGATCTTCAATAGCAGCGGCTGCGGAGATTTT 8999
Qy 5616 TCGCTGATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 5675
Dh 9000 TCGCTGATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 9059
Qy 5676 AGTTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 5735
Dh 9060 AGCTTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 9119
Qy 5736 GAAAGGCGCGCTGCGGATGCGCATTTTCTTATCTGAGAGAGCATATATGCGGCT 5795
Dh 9120 GAGGCGCGGAGCTGCGGAGCGGATTTTGAATCTTGAAGAGGCTATATGCGGCT 9179
Qy 5796 CTGCGAGCAGATTAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 5855
Dh 9180 CTGCGATGCGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 9239
Qy 5856 CAGGCGATGAGGCTGAGGCTTCTGCTGATGATGATGATGATGATGATGATGATGAT 5915
Dh 9240 GATGCGATCTGAGATGAGGCTTCTGCTGATGATGATGATGATGATGATGATGATGAT 9299
Qy 5916 CCGGCGCATTAACCGGATGCGGATGCGCATTTATCTTATGATGATGATGATGATGAT 5975
Dh 9300 CGCGCGGAGCAACCGGAGCAGGATGCGCATTTATGCGGATGATGATGATGATGATGAT 9359
Qy 5976 CCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTG 6035
Dh 9360 CCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTG 9419
Qy 6036 GAGGATCTGAGGAGTGAACCAACCGGCTGATGAGGAGTGAACCAACCGGCTGAG 6095
Dh 9420 GAGGATCTG--CAATGAGGAGTGAACCGGCTGATGAGGAGTGAACCAACCGGCTGAG 9476
Qy 6096 AGGCTCTCAAAAGTTTTCGCGGAGCTTACCGGCTTATTTCTTCAACCGGCTGAG 6155
Dh 9477 AGGCTCTCAAAAGTTTTCGCGGAGCTTACCGGCTTATTTCTTCAACCGGCTGAG 9536
Qy 6156 CGCTGATGCTGATGAGCTTGTGCTGATGATGATGATGATGATGATGATGATGATGAT 6215
Dh 9537 TGTGATGCTGATGAGCTTGTGCTGATGATGATGATGATGATGATGATGATGATGAT 9596
Qy 6216 GCGAAGGCGGAGCAGGAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 6275
Dh 9597 GCTTTCGAGAGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 9656
Qy 6276 TCGAAGCCGCGCGCTTACAGCGGCTGCGGAGTGAAGTGAAGTGAAGTGAAGTGAAG 6335
Dh 9657 CGCTGAGCCCTGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 9716
Qy 6336 AGGAGTGGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 6395
Dh 9717 AGGAGTGGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 9776
Qy 6396 TTTGAGTGAATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 6455
Dh 9777 TTTGAGTGAATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 9836
Qy 6456 ATACGCTGCGCGGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 6515
Dh 9837 ATACGCTGCGCGGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 9896
Qy 6516 GCGTACTGATCAAGCGCTGATTTTGGCTGCGGCTTACGCTTACCAATATGCAAGGAG 6575
Dh 9897 GGGTGTGATTCGCGGCTGAGTCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 9956


```
QY 3523 CAGCGGCGCGGCAATCAATCTGGTCTGTCATGCGGAAGACGTACAGGAAGCCAGTTT 3582
Db 245 CAGCGGCGCGGCAATCAATCTGGTCTGTCATGCAACAGATGATTTGATGAAAGCCAACT 304
QY 3583 CGCTGGCTCGAGCGCGCTCTTCCGCGCGCTGAGCCCGGTTATCAGTAAGCTTCCCAAC 3642
Db 305 CGTTGATAGTCCGCTGGTGGTTCATCATCTGCGCCGACATACAGTAAGCTTCCCAAC 364
QY 3643 CTGCGTGGCAGCTGATGATGTAATATTGCTCATGCTGCTGAGAGATTTTCCCGGAC 3702
Db 365 CGCGGTGCTAAGTGAACAGCGGCTACTTTGTAATCTTCTCAGCGTTTGGCTGAGGTT 424
QY 3703 TTACAGAGGTGCTGGTGGCGCGCTACGCAACCGAGCGCGCTGACAGAGGTCTCACCC 3762
Db 425 TTACAGAGCAAGTTTGGCGCGCACTTGTGATGATACCGCGCTGCGAGAGTTAATGCG 484
QY 3763 ACCGGGCTCAGACTGCGCGATGCGCGGATGTTACAGCGCGAGCGGATGACGAGCGC 3822
Db 485 GAATCTGTTCGTTGAAAAGGTCAGGTTATCGGTGCGCGCGCTGATTTGACGGGCGG 544
QY 3823 GGGCTGCAAGCGACACCGCATCTGCAAGCTCGGCTATACGCAATTTTCTGCTCAGAGTG 3882
Db 545 GGTATGCGCAAAATTCAGCACTGACGTGGGCTTCCAGCGTTTATTTGCGCAGGAAATG 604
QY 3883 CAATGCGCGCGCGCATGCGCTGACAGCAATATTGATGAGCGCGAGGTGATCAG 3942
Db 605 CGATTGAGCAACCGCGCATGTTTATCTCTCCATATCATGATGACCGAGGTGATCAG 664
QY 3943 CAGAGGGTTATCGCTTTGTTTACACCTGCGCTCAGTGCAGCGCTTTTACGATTTAA 4002
Db 665 CAATATGGTTATCGCTTGTGTACAGCTGCGCTGCGCGACCAAGTTGTAAATTTGAA 724
QY 4003 GATACCCATCATCAACCATATGCCACGCTGATGCGGACAGGCGCGCTGACATTAAC 4062
Db 725 GACACGCACTATATGTAATGACACATTAATCTGTAATGCGCGCGCAAAATTTTGC 784
QY 4063 GATTATGCCCAACGCGCGCTGGAATTTGCGCAGCTGCTGCGGAGAGACAGGCTCG 4122
Db 785 GACTATGCGCGCAACAGGTTGGAGCTTACAGACCTGCTGCGAGAGAACAGGCGCC 844
QY 4123 CTGCGCATCAGCTCAGCGCGGATATCGATCATGTTCTGCAACAGCAGCGGCAACCG 4182
Db 845 TTACCATTTACTGTGTGGGCAATGCGGACGCAATTTCTGG--CAGCAGCGCGCCCTGCGC 901
QY 4183 TGCAGCGGCTGCGCGCGGACTGTTTACGCGCACACCGGTTACTGCTGCGCGCGC 4242
Db 902 TGTATGATTAATGATGCGGCTGTCTGTTCTTACACACCGGCTATTTCACTGCGCGCG 961
QY 4243 GTGCGCTGCGGGAAGATTTGCCAGCACTGCTGCCCGCGCAGCTCACACGCTGAGCAC 4302
Db 962 GTTGCCTGCGCGAGCGGCTGAGTGCATTTATGCTGCGGCTCAATTCACCAT 1021
QY 4303 TGCATCGAATCTTTGCGCGTACAGCATGCGCGGAGCAGCGCTTTTCCGTCTGTAAT 4362
Db 1022 GCCATTAACGATTTTGGCGCGAGGCTGCGAGCAGCAGGCGCTTTTCCGATCTGAT 1081
QY 4363 CGCATGCTGTTTCTTCCGCGAGCGCTGAAACAGGCTGCGCGTAAATGACGCTTTTAC 4422
Db 1082 CGCATGCTGTTTAAAGCGGACCGCGCATTCAGCGTGGCGGTTATGACACGTTTAT 1141
QY 4423 CGGCTTGAAGCGCGATGATTAGCGCTTTTACGCGGGAACACTGCGCTCAGCGATPAA 4482
Db 1142 GGTTTACTGGAAGATTTATTTGCGGTTTATGCGGAAAACATCAGCTTACGATCGG 1201
QY 4483 GCAAGCATTTCTGTGCGGCAACCGCGGTCCTCTCGCGGAAGCGCTGCGGCAATTGAG 4542
Db 1202 CTACGTATTTGAGCGGCAACCGCGCTGTTCCGATATTAGCAGATTGCAAGCATTAATG 1261
QY 4543 ATGACCTTCCGTTAACAGGGAATAATGAAACGACCTAATGATTTGGCGAGGCT 4602
Db 1262 ACGACTATCTGTTAAAGGACACTA--CATGAAACCAACTACGATTAATTTGGTCAGGCT 1318
QY 4603 TCGGTGCGCTGCGCGTGGCGATTTGTTCTGCAAGCGCGCGGCTGCTGCTGAG 4662
Db 1319 TCGGTGCGCTGCGACTGCGCAATTCGTTTACAACTGCGGGGATCCCGTCTTACTGCTTG 1378
QY 4663 AACAGCGATTAAGCTTGGCGGAGCGCGCTATGATGATACAGATTCAGAGTTTAACTTTG 4722
Db 1379 AACAACTGATTAACCGCGCGCTGCGGCTTATGCTACAGAGATACAGGGGTTTAACTTTG 1438
QY 4723 ATGCGGCTCCGAGGCTATTAACGATCCGAGCGCTATCCAGAGCGGCTTTAACGTCGAG 4782
Db 1439 ATGAGGCGCGGAGTTATACAGATCCAGTGCATTAAGAACTGTTTGCAGTGGCAG 1498
QY 4783 GCAAGCACTCAGATTAATGATGCACTGATGCGGAGCGGATTTTATGCTGCTGCT 4842
Db 1499 GAAACAGTTTAAAGATATGCAACTGCTGCGGTTAGCGGCTTTTACCGCTGTGTT 1558
QY 4843 GGGAAAGCGCAGGCACTGCACTACGACAAATACAGGCGAGCTGAGACAGATTTG 4902
Db 1559 GGGAGTCAGGGAAGTCTTTAATTAAGATACATCAAAACCGGCTGGAAGCGCAGATTC 1618
QY 4903 CCACTTTTATCCCGAGAGATGTCGCGGTTACCGGCACTTTTGGCTTATTCAGAGATG 4962
Db 1619 AGCAGTTTATCCCGGAGTGCAGAGTTATGCTCACTTTGAGACTATTCAGCGCG 1678
QY 4963 TGTTCGTAAGGCTATCTGAAAATGAGGACCGTACTTTTCTGCAATTTCCGCACTATG 5022
Db 1679 TGTTTAAAGAGGCTATCTAAAGCTCGTACTGCTCTTTTATGCTTACAGAGACATGC 1738
QY 5023 TCGGTGCGGCGCAACAGCTGGGTGCGCTGACAGGCTGCGCAGTGTCTACAGCATGCTG 5082
Db 1739 TTGCGCGCGCACCTCAACTGCGCAAACTGACAGCATGAAAGCGTTTACAGTAAGGTTG 1798
QY 5083 CGAAATTTATTCANAGATCATCTGCGCAGGCTTTTCTTCACTGCTGCTGCTG 5142
Db 1799 CGAGTTCATCGAAGATGAACTGCGCAGGCGTTTCTTCTCACTGCTGCTGCTG 1858
QY 5143 GCGGTAATCTTTTGCACAGCTCTTCTCATCTTAACTTATTCACGACTGAGGCGCAAT 5202
Db 1859 GCGGCAATCCCTTTCGCACTCATCTTATACGTTGATACAGCGCTGAGAGTGAAT 1918
QY 5203 GGGGCGTGTGTTTCCGCGCGGAGTACCGGTGCGCTGTTATGCTGATGCGCGGCTGT 5262
Db 1919 GGGGCGTGTGTTTCCGCGCGGAGTACCGGCTCATTAATGTTACAGGAGTATTAAGCTGT 1978
QY 5263 TTGCGATTTGGGCGGTGAATGCTGCTCAACCGCAAGTCAGCGCTGGAAGCGAGG 5322
Db 1979 TTCAAGATCTGGGTGGGAGTGTGTTTAAACGCAAGTCAGCATATGAAACGACAG 2038
QY 5323 GTAACCGCATCAGCGGTGTCAGCTGAAGAGTGGCGCGCTTTTGCAGCGCGCGCTTG 5382
Db 2039 GAAACAAAGATTTGAAGCGGTGCAATTTAGAGGACGCTGCAAGTTCAGCAAGCGCTG 2098
QY 5383 GGTCAATGCTGACGTGTGCAATCTTACGATGCGCTGTTTAAAGCAGCATCTGCGGCGC 5442
Db 2099 GGTCAATGCAATGATGTGTTTCAATCTTATGCGACCTGTTTAAAGCAGCATCTGCGGCGC 2158
QY 5443 GTAAGCGCGGGAAGCGCTGAAGCGCAAGGAGTGAAGCACTGCGTGTGTAAT 5502
Db 2159 TTAAGCACTCAACAACTGCACTGAAGGCAATGATCTCTGTTGTGCTTAAT 2218
QY 5503 TTGCTCTTAATCAATGCCACCGGAGCTGCGGACCAACAGCTGTGCTTGTGCTGCGCT 5562
Db 2219 TTGCTTTGAATCAACATCATGATCACTGCGGATCAACAGGTTTGTGCGCGCGGCTT 2278
QY 5563 ATGCTGAATGATGATGATGATCTTCAATACAGCAGCAGCTGCGGGAAGATTTCTGCTGT 5622
Db 2279 ACCGCGAGCTGATGATGCAAAATTTTATCATGATGCTTCCGAGAGGACTTCTCATTTT 2338
QY 5623 ATCTGATGAGCGCTGCTGCAAGGATCCGTGCGTGCAGCAGCGGCGGTGCGGAGTTT 5682
Db 2339 ATCTGACGCGCGCTGCTGCAAGGATCTGCACTGCGGCTGAAAGTTGCGGAGTTACT 2398
QY 5683 ACGTGTGCGCGGCTGCGCATCTCGTATCCGCGCAATTTGACTGCAACAGGAAGGC 5742
```

[illegible]

Oy	6823	AGTGCACATGACCGCGTCCCGGATCGGATATGGATATACCGCGACGACACGTAAGAGCGGAAAA	68882
Oy	6823	AGTGCACATGACCGCGTCCCGGATCGGATATGGATATACCGCGACGACACGTAAGAGCGGAAAA	68882
Db	3336	AGTTGAACAGGCGCGGTACGACAGCCTTGGATCAGCGGCAATCAACGACACGCGCGAAAA	33595
Oy	6883	GCTGGCGCTGTGTGTGAAAGATGCGCGCGCTGTCTTACTTTCGGCTTGCTCTCCGA	69424
Db	3596	ATTAAAGCTGTGTGTGGCCGCTCTGTGTACAGGCTTACTTCCGGATATGCGGCGCTCATTC	36555
Oy	6943	GGCCCGTCTGCGCGGTCTGTGGCAGCGCTCCGCGTTGACACAGCGCCATGGCGCTGGCGC	70022
Db	3656	TCCCGCGCGCTGCGCATCTCTGGACGCGCCCG-----CTTAAAGCCATGCTTTCGCG	37099
Oy	7003	AGCGTCGCTGCGACCTTGTCGCAACGATGGGCGGTAAAGAAAGCAAGAAAAAGCGACCT	70622
Db	3710	AGCGTCGCTGTAAGTTTGTGACAGGGCGCGCGCATATGAGAAAGCAAGAAAAAGCAACCT	37658
Oy	7063	TCCCGTCCCGCACCGCATGATGTCATCGGTGCGCCATATACAAACGCTTCAGATAGCTT	71222
Db	3770	TCTTTGCCCTTGACGCGGTATGATATACGATGCGCCATATACACCGTTTGAAGTAGCC	38299
Oy	7123	TTTGCTGGAATATATAGCGAAACGCGCAGCGTTGATGACACAGCGCGCTCATGACATATAA	71822
Db	3830	TTTGCTGGAATATATAGCGAAATGGCCAACTTGAATGACACAGCGCGTGTGACATATAA	38899
Oy	7183	TAGAGAGCGGCATAGTATGCTATGCTATGCTCAGCGCCATTCATCTGACGCGACGCGCTGT	72424
Db	3890	TAGAGTATATCATATGCGCGCTATCTCTGCGCCATTCATCTGAGCGCGCATATTCCTGT	39499
Oy	7243	ACACCCAGCCAAATCAGGCGCAATGACACACGCGCAACACACGCGCATACAGGTCGTT	73022
Db	3950	CTGCCAGATATTAACACAGAGATGATATGACAAACAAACACGCGCATATAAGATCTT	40099
Oy	7303	AGCTCAATTTGCTCTCATGTGTTCATATGATGCAAAATGCGCGCCCATCCGACGCA	73622
Db	4010	ACTTCAACGACCTTTACGCGTTCATGATGTAAGATGCAATCCCAACCCGACGCG	40659
Oy	7363	TGCAATATATTTATTCGACAGCGCGCGCGACATCTCCATCACTATTCACGCGTACGCG	74222
Db	4070	TGCAATATATTTGTGTGTGCGACAGTGCAGCAATCACTTCCATGCCATACGCTAACGAA	41229
Oy	7423	AGGATTAAGCGCATTCATTAACGACAGATCATGTGTCATTTGCGAAAGATGAAGTA	7480
Db	4130	ACGATCAAGGCGATTCCAATTCACAAATATATTTCTCCGGATGAGAGCGTTCGACGA	4187
RESULT 8			
BX571873			
LOCUS	349107 bp	DNA	linear BCT 26-SEP-2003
DEFINITION	Photorhabdus luminescens subsp. laumondii TTO1 complete genome;		
ACCESSION	BX571873 BX470251		
VERSION	BX571873.1 GI:36787441		
KEYWORDS	complete genome.		
SOURCE	Photorhabdus luminescens subsp. laumondii TTO1		
ORGANISM	Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales; Enterobacteriaceae; Photorhabdus.		
REFERENCE			
AUTHORS	1 Duchaud,E., Rusniok,C., Frangeul,L., Buchrieser,C., Taourit,S., Boos,S., Bouraux-Bude,C., Chandler,M., Dassa,E., Derose,R., Derzelle,S., Freysinet,G., Gaudialtic,S., Glavaudan,A., Glaser,P., Medigue,C., Lanois,A., Powell,K., Signier,P., Wingate,V., Zouine,M., Boemare,N., Danchin,A. and Kunst,F.		
TITLE	Complete genome sequence of the entomopathogenic bacterium Photorhabdus luminescens		
JOURNAL	Nat. Biotechnol. 11 (1) (2003) In press		
AUTHORS	Duchaud,E., Frangeul,L., Rusniok,C. and Kunst,F.		
TITLE	Direct Submission		
JOURNAL	Submitted (23-APR-2003) L. Frangeul, Institut Pasteur, Genopole, 25 rue du Docteur Roux, 75724 Paris Cedex 15, FRANCE. E-mail: litangeu@pasteur.fr		

```

FEATURES
source
1. 349107
/organism="Photobacterium luminescens subsp. laumondii TT01"
/mol_type="genomic DNA"
/strain="TT01"
/db_xref="taxon:243265"
complement(80..3225)
/gene="tccC1"
/locus_tag="plu4167"
complement(80..3211)
/gene="tccC1"
/locus_tag="plu4167"
/function="unknown, probable insecticidal toxin"
/note="Highly similar to insecticidal toxin complex protein Tccc of W14."
/codon_start=1
/transl_table=11
/product="insecticidal toxin complex protein Tccc1"
/protein_id="CAE16539.1"
/db_xref="GI:36787442"
/translation="MSTPDTALVTQTPVSVLNDNRGLSIDRIGFHRVIGSDITRYT
RHQYDARYLNHSIDPRLYDAKQADNSYKPFVQYDLAALMTESYDAKRTYALND
IEGRVMTWNAIVGRQTRRYGNTLPGRLSVSQVDPQETSQVTERPIVMENTTAK
EYNLSGLCIRHYDAGVTRIMSQSLAGLILSQSHQWLAEGQANMSGDETVQRMILA
SEVYTOSTTNAIGALLTQTDKAGNIQRLAYDIAGOLKGSWLVKQGEQIVKSLSM
SAAGHLEEHNGVITTEYSYEPETORLMGTTTAAEESQSEARVLDLRYEYVPGN
VISIHNDAEATRFWRNOKVEBENRYVYDLSLYQMSATGRERMANICQOSNOLPSAMPV
PADSTYTYNORYTSYDRGNTLVQIRHSSPATQNSYTTENTVSRSNRVTLSITTPD
TQVGLFPAGHQKTLTPGQNLDMNMRKELRVTVSENSDSDEWYRSDDGRLLK
VSEQQTNSYQVQVYVTLPGLELRITGVADTQNTLVITVGEAGVRAQVRLHSESK
PADIDNNQVRSYNNLGSQIELDSEGOILQSEYYPGGTAVMAAKQTEASFKYI
RYSGERADATLQYGYRYQYQVGRMLSDAPAGVQGLANVRVWNNPITLTDHGL
ASPNNRNTFMFAPSLPRKPDMSAMBRGOKIRAGLIGLIGLAAIAGTASA
ALPILIGAAVAGAGIGALMGVNSLLEKSGALLARLVQGSITVQSAAGAAASAA
AAYGARQGVAVASAGAVTGAVSMMINADRGIGAGIAGSAGTITDMLGISTSLT
HEVAAAGAGAGMIITGQSTRAIGIAGIGTYGSMIGFGLDVAISPAGLANVAG
YAGIGEMAINRYIGGFGFLRGLGRVSPYAGIARQLVAFSIVARVFETIPSLVGS
LAGIGTGHVWRGRDWSVRGLSAGSGIDLHAGMIGNQIRGRVLTGTGIANADYG
TSAIGARVIVSL"
complement(3220..3225)
/gene="tccC1"
/locus_tag="plu4167"
complement(3345..8051)
/gene="tccB1"
/locus_tag="plu4168"
complement(3345..8039)
/gene="tccB1"
/locus_tag="plu4168"
/function="unknown, probable insecticidal toxin"
/note="Highly similar to insecticidal toxin complex protein Tccb of W14 Putative transmembrane protein."
/codon_start=1
/transl_table=11
/product="insecticidal toxin complex protein Tccb1"
/protein_id="CAE16540.1"
/db_xref="GI:36787443"
/translation="MISTMEKQINESORDALVTGYMNFVASTLKGVDGQPVYEDLYE
YLIDPREVADSEVTSVAQALASIQQWTRLVNGSEPRGQAMELSTANERDNDQYA
IMAGAERYNAENYISPIITROKSHFSELETTLNQRDLPPDYQDAVIGLYNEFA
VSNLYVLSGYINDKFDQALYFLGRITTKRYRYWROMDISKRDPAQNPVNPVNC
NDWQBITLPLSGYDVLVETVAPVFNDRLVYAWWRBPVAVQKDAGKNIKGTAYSIK
FGYKRYDDTWAPNTTLMTEQADESSTORSILIGESKTLAKNMLATDERSIDP
TEETDSNRYGLMGVFNRTGTPHQNRPAVGYLXCDSAFNNHILRLPSKDELSFT
YRDETNDONCQFPAVDKXYITKVISGKPEDEENMGCKVDLKLGTGAYVYD
QDGTIHLQTTNKDPTNRHIFGYNIVYDSKSGYGTWGSNEFTYLDHGNYYTH
NAALINYVAGYGGSVNGWTALRQVRNVEGALISPLLDHHTITVQSYLAWEGETA
SYLPIPDGTMLDMSKINPALSINKLSEVTSIPDPTLTIKNPSKIDANRFYOE
INATLADGRNLFKHYSEQTFQASGATPYTTLSEADFSIDPDKNVLQVCLNVMYMH
YDRSGRKGAVSMKWNVVELDQSKAPAPILRISYDSKRGLOVYDGTSSIP
AKRILNTFVPTLIEKANIGLDSLDTLQADPSLADIVTDDGSEPMDDGSGNGYF
WELEFPHLPELVATPFANEQSPAKSKLHITFDPMKGRPHNAPAYNNRPLVEYSD
LSRLGDPIDDTQAVAPVLYQAVSIAYVSIAGQDMWRQJLTDGGLQAVVYN
LAAELGPRPDVSLSTWTPTQTLDTLLAAGOKTALRDEHOLANSNTVLPALPGNNISH

```

```

RBS
/locus_tag="plu4169"
complement(8134..11034)
/gene="tccA1"
/locus_tag="plu4169"
complement(8134..11034)
/gene="tccA1"
/locus_tag="plu4169"
/function="unknown, probable insecticidal toxin"
/note="Highly similar to insecticidal toxin complex protein Tcca of W14."
/codon_start=1
/transl_table=11
/product="insecticidal toxin complex protein Tcca1"
/protein_id="CAE16541.1"
/db_xref="GI:36787444"
/translation="MNQVAPSLSRTEEIHNLPFKLTDLGYTSFVDVMPRRPFIRE
HRADLGASAKMVDLAVYAGVQVLFHRRNSLEAVQFSIRSPFSVGPYQAQFLDA
NKTMEVAPSGSEANDAPAVATLHLYQLDDEKSGASIMNTLVERRRDIALLIN
DKINVEYIPLOLVNEILSKAIQKLSITLBEVNAKSLSTRYPNNLPYHIGQOQT
AGSVLTGLQDITLPTQDLIPNMFATPKKSDADARALTRLOIMASQISPEQKID
VEYVQGFYLANGDSLITNANSDMTVLTDRSLVYQVLEMLCTYGSSTYKSDM
VSGDVTATPFATGARITAGKPOATLTSNGAEARHALVNNLTDDKDIRINRYRL
QKMLNIPYEDIDILVTSAMDATGNTALSNNDTLMLEGVFKYQAKYGTARQFQGM
LRVATPEALPATPFIDQVNSYGTPTPEVDNQPFVYLTGCGGARVKAHISTALG
LNHRQFLADNIAROOGNITOSTLNCNLFVSAFYRLANLARTLGNPSPFALVDR
LDAGTGVQOOLAGKPIIMVPOKSDPLADILSLQALSVIAQMOOORDPIETLLL
LSSGFTTQAGTDDQNFIRQVQNGNSTVSKTILTSRGAPEVDTNRGAIDMPALLS
AGNSPLIDQVLTQDAGIESVITVAVRQSLSEDEKCLATITLNTLYQAOQOQGA
VSLIAQTLNQSQSLPALLRMSQTYTQWLSATWALKDVTYKADIAPAGLSLREYV
RRSLIQOQFTLSPAMVQTLIDYPAYGASAEVVTDISLMWLYLSRSHYDILQVKG
GGTEVDYALVLRVANAATPLSODAAQTALTLIGMETNELQAAWEVYGIATKTPQD
TLRLQOASQGTGAVTQOQGYVLSRSDSYTLMOSTGQAMVAGSHVKGSH"
complement(11462..11943)
/locus_tag="plu4170"
complement(11462..11929)
/locus_tag="plu4170"
/function="unknown"
/note="unnamed protein product; Similar to pyocin S3 immunity protein and to Photobacterium luminescens proteins"
/codon_start=1
/transl_table=11
/product="insecticidal toxin complex protein TccB1"
/protein_id="CAE16542.1"
/db_xref="GI:36787445"
/translation="MAINIEALINSKTYQOEIFDEGLIPYKTPSGYPGDSIDSLDM
AKGRTLSRREKXVLEITLPIRNKRPFIEPNKLPSGLPIEMRKLIHQHGRGE
KSLPFRKRLTXEIGWTELYTLDLFRIPTSQVDYDUQERRVLTFLPTSEIRN"
complement(11938..11943)
/locus_tag="plu4170"
complement(12038..12505)
/locus_tag="plu4171"
complement(12038..12505)
/locus_tag="plu4171"
/function="unknown"
/note="unnamed protein product; Similar to pyocin S3 immunity protein and to Photobacterium luminescens proteins"
/codon_start=1
/transl_table=11
/product="insecticidal toxin complex protein TccC1"
/protein_id="CAE16543.1"
/db_xref="GI:36787446"
/translation="MTVNVDAALINSKGSQVQEIFDEGLIPYKTPGTGPGASPIALNM

```

```
AKGMLAPKRDKII.PAVEL.FILDOKRPI.YOPFNEI.PSLKRI.MTRVWVHROFKRBE
KALPRKFLKDYVWTELYTLDFRIPTSMQVYDILBOVKSAVFLSTSEVLM"
gene
  complement (12597..13081)
  /locus_tag="Plu4172"
  /locus_tag="Plu4172"
  /locus_tag="Plu4172"
  /locus_tag="Plu4172"
  /function="Unknown"
  /note="unnamed protein product; Similar to pyocin S3
  immunity protein and to Photobadus luminescens proteins"
  /codon_start=1
  /transl_table=11
  /protein_id="CAE16544.1"
  /db_xref="GI:36787447"
RBS
  /translation="MIDVETIINSLEKTDMLVENKRIPEKREPEEGYSRSVEBE
  TGLIVDIIIVLISVQFTLVNTVDITGYSDMDPAFLHKNKAIVRAPWGEPPYIS
  GSEPIRPIVIGLVPFDPAYTRKLANDOYPTNREIRFLYQDILFVDALEFOSVNDN"
  /complement (13076..13081)
  /locus_tag="Plu4172"
  /locus_tag="Plu4172"
  /complement (13174..13652)
  /locus_tag="Plu4173"
  /complement (13174..13635)
  /locus_tag="Plu4173"
  /function="Unknown"
  /note="unnamed protein product; Some similarities with
  pyocin S3 immunity protein and similar to Photobadus
  luminescens proteins"
  /codon_start=1
  /transl_table=11
  /protein_id="CAE16545.1"
Query Match      15.5%; Score 1345.8; DB 1; Length 349107;
Best Local Similarity 61.5%; Pred. No. 1.3e-242;
Matches 2191; Conservative 0; Mismatches 1367; Indels 5; Gaps 2;

QY 3421 GATGCAATTCGTGCGTGGTGGCGCTGCTAAAGCGGCTGATTCGCTGCTTTACGCGCAG 3480
DB 202949 GATCTGATTCGTGCGTGGTGGCGCTGCTAAAGCGGCTGATTCGCTGCTTTCCGCGCAG 203008
QY 3481 CTGCGAGCGCGCACTTAAAGTTTGTCTAATGAGAGTCAGCGCGCGCGCGCAATCAT 3540
DB 203009 TGTAAACACATCTGCGAGGTGTGCTTATGAAAAACAGAAACAATAGAGGCAATCAC 203068
QY 3541 ACCTGCTGCTTCATCGCGAAGACGTACAGGAAGCGGATTTGGCTGCTCGAGCGCTG 3600
DB 203069 ACCTGCTGCTTCATCGCGAAGACGTACAGGAAGCGGATTTGGCTGCTCGAGCGCTG 203128
QY 3601 CTTTCGCGCGCTGCGCGGCTTATCAGGTACGCTTCCCGACCGCTGCGCGCGAGT 3660
DB 203129 ATTACCTATCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 203188
QY 3661 GGTGAATATTGCTGATTCGCTGCGAGATTTTCCCGGCACTTACAGCAGGTGCTGCT 3720
DB 203189 CATTCATATTTCAGTATCATCATCCCAACCTTTCAGAGCATTCATTCATTTTGGGC 203248
QY 3721 GCGCGCTAGCAGCAGCGCGCGCTGCGAGAGTCTACCCACCGGGGTCAAGTGGCG 3780
DB 203249 GAACTGATACAGACTGCTTATTTATTTGTAACAGAGCTGACTCCACAGAAAGTTTACTTACAG 203308
QY 3781 GATGCGCGGATTTTACAGCGCGCGGCTGATTTGACGCGCGGCTGCGCGCGCAACCG 3840
DB 203309 GAGCGCTGCTCTTAAGTGTGCGCGCGAGTATTTGATGAGAGAGCTGGCCACAGAGCA 203368
QY 3841 CATTCGAGCTGCGCTATCAGGCAATTTGTCGCTCAGAGTGGCAACTGCGCGCGCGCAT 3900
DB 203369 TTTATAGAGAGTGGACCCAGGCAATTTTTCGTCAGGAATGGAGCTGGAAGAGTGGCAC 203428
QY 3901 GCGCTGAGAGGCAATATGATGAGACGAGCGTGCATACAGACGAGGCTTATTCGCTT 3960
DB 203429 TCTTTAACCAACCGGATTTTAAATGATATAGTGTGGACAGGATTAAGGATTTT 203488
QY 3961 GTTTACACCTGCGCTCAGTGCAGCGCTTATCTGATTAAGTAACTCCATCATCAAC 4020
DB 203489 ATCTATGCTCTGCGCTTCTCTCTCACTGCTGCTGATTAAGAGCACTCATTTACGTTGAT 203548
```

QY	50959	GGATCATCTGGCGCAGGCGCTTTTCCCTTACACGTGTGGCGGGGTAAACCTTTGC	5158
Dp	204629	TGAGACCTCGCGCCAGGCAATTTTCATTCACTCGCTATTGATAGGAGGTAAACCTTTGC	2046888
QY	5159	AACGCTTTCGATCTATACCTTAATTCACGCACCTGAGCGCGAATGGGGCGTGTGTTTC	5218
Dp	204689	TGCATCTTCTATCTATCTATCTTTGATCCACGCACCTGGAACGAAATGGGGAGTGTGTTTC	204748
QY	5219	GCGCGCGGTAAACCGGTGCGCTGTGTGATGTGCATATGGCGCGCTTTTCGGATTTTGGCGG	5278
Dp	204749	CGAGAGAGAAACACGCGCTCTGTGTGAGCAATGTGAAATGTTTATCTGACATTTGTGTG	204808
QY	5279	TGAATCTGCTCAACGCGCGAAGTCAGCCAGCTGGAGACCGAGGTTAAACGCATCAACG	5338
Dp	204809	AGAAATTGAGCTTTAATGCAAAAGTAAACGCTTCAACACCCATGGAATCGGCTCAACG	204868
QY	5339	TGTCACGCTGAAGAATGGCGCCGTTTTCGCCGCCGCTTGCGTCAATGTCTACGT	5398
Dp	204869	TGTCCAATTAGCTGATGAGCAAAATATGACATGTACGTTGTGGCCCTCAAAATGCCATGT	204928
QY	5399	GGTGCAATCTAGATTCGCTGTGTTAAGCAGACATCTCGCGCGCGTAAACGCGCGCAAC	5458
Dp	204929	TATTCATACCTTAATAGCACCTGTAGTACGATCCGCTCGAATATGCCCGCTGAAAC	204988
QY	5459	GCTGAAGCGCAAGCGGATGAGCAACTCGCTGTTTACTCTATTTTGTCTTAATCATGC	5518
Dp	204989	ACTGTGCGTAAACCAATAGCAACTCACTGTTCGTCTTAATTTTGGCTCGAACATCA	205048
QY	5519	CCACCGGAGCTGGGGGACACACAGGTGTGCTTTGGTCCGCGCTATCGGAATTGATCGA	5578
Dp	205049	CCATAGCAATTAGCCCATCAATCAATCAATTTGCTTTGTGTCCACGCTAATAGGAACTTATCGA	205108
QY	5579	TGAGATCTTCATATAGCAGCAGACTGGCGGAAGATTTCTCGCTGATCTGATAGCGCCTG	5638
Dp	205109	AGATATTTTATTCATATGACCGACTGTCAAGAGATTTTTCACCTCATCTCCATGCCCTTC	205168
QY	5639	CTCCAGCGAATCCGTGCTGTGGCACCGGCGGGCTGTGGCAATTTTACTGTGTGGCGCGGT	5698
Dp	205169	TGTTTCTGACCCATCCCTGTGCTCCCGAGAAATGCGAAATTACTAAGTTCTTGACCTGTG	205228
QY	5699	GCGCGATCTCGGTACCGCGCGCAATTGATCTGGGCAACAGAAAGGCGCGCTGGCGCATCG	5758
Dp	205229	ACCAATTTGGGGGACAGCTAATTTAATCTGGGATATAGAGGGCCACGCTTACGTGACCG	205288
QY	5759	CATTTTGTCTATCTGAGAGCACTAATGCCGGGTCTGTGCACAGCAGTTAGTACACA	5818
Dp	205289	TATTTTCCGTTACTGGAATAATCTATATATGCGCGCTGTCTAACAAGTATGTCTGTCA	205348
QY	5819	CCGTAATGTTTACGCGGTTTGAATTTTTCGGGACACGCTGCACGCGCATCAGGGCTCAACGTT	5878
Dp	205349	TGTAATTTTATCCCAATTTGATTTTGTGTACCAACTCAATGCCAATCTCGGTTGGGCTTT	205408
QY	5879	TTGCGTCGAACCCATTTTGTAGCGAAACGCGCTGTGTTCCGGCGCGCATTAACCGCGATGCCGA	5938
Dp	205409	CTCCTTTGAACCGTGTGTGCTACCAAAAGCCTTGTTGTTCCAGCACACAATTCGGGACAACCG	205468
QY	5939	CATTACTAACCTTATCTGTGTGGGGGTGGACGCTATCCCGGTCCGGTGTGTGCAAGCGGT	5998
Dp	205469	CATTATATCTGTATCTGTGTGGCGGAGACGACGATCTGTGAGCAGGTATTTCCGGCGGT	205528
QY	5999	GATCGGCTCCGCGAAAGCACCGCCAGCTGATGTGTGAGAGATCTACCGGATGAACCAA	6058
Dp	205529	GATTGTGATAGCAAAAGCACCGGCCACATTTGATGTAGAGGATATGCTTAATATGA---AT	205585
QY	6059	CCGCCGCTGATTTAGCAGGTCAACGCAACCTATGGCGCAGGGCTCCAAAAGTTTGGCCAGC	6118
Dp	205586	CCTGGCTACTTAACACGTTTACTCAGATTAATGGAACAAGGCTCAAAAAGTTTGGCAGT	205645
QY	6119	GCTACCCGGCTTTTGGATCTTCAACGCGCGGAGTACGCTGATGTGTGTAACGCTGATGTG	6178
Dp	205646	GTCACCCGACTTTTGGATTCAGACACACACACACAGCAATGATGTGTGATGCGCTGTGTGC	205705
QY	6179	CGTACGTGTGACGATGTGTATAGTGTCTGACGCTGTGGCGAAAGCGGACGCGACGCGC	6238

Db	205706	CGTATTGTGATGATATATATATAGACGGGCAAGAGCTGGGAGACAAATAAGCACTGTTGAT	2057655
QY	6239	GTGGCGGATGCACAGCGCGGATGCGCCACTGCAAAATGAAACCCGCGCGCTTACAGC	6238
Db	205786	AAGTATACGCGCCCGGAGAAACTTCAATGCTGCAATACCTGACAAAGACAGGCTTACGAC	2058235
QY	6239	GGTGCCCAATGATGATGAACCAAGCGTTTGCTGTCCTTTACAGAAAGTGCGCGCTGACGCAATCAG	6238
Db	205826	GGTTTGGCGATGACCGAACCCTGCTTTTGACAGCTTTTCAGACAGTACCTTGAATATGA	2058855
QY	6359	CTTCCCGACAGCTGGCTTTTGATCATCTGAAAGGTTTGCATGATGATCCGCGTGAAGA	6418
Db	205886	ATCCCTCAGCAACAGCGCTTCGAAACATCTGAAAGATTTGCAATGGATGATCTCTGTGAG	2059454
QY	6419	CGTATGCGTGTTTTGGGGACACGCTGCGTTACTCTGATACAGTGGCGCGCGTGTGGG	6478
Db	205946	CTTATCAGAACGTTGATGATATCTGTGAAATTTGCTACCAATGTCGGGAGATGTTGG	2060050
QY	6479	TTAATGATGCGCGCGGATGAGGCGCTAGCGTGAATGAGCGGTAATCGATCAGCGCTGTGAT	6538
Db	206006	CTGATGATGCGAAGGTTATGGGGGTTTGCGAGAGGAGTGTCTGGATCGGGCGTTCGAT	2060655
QY	6539	TTGGGTCTGGCGCTTTCAGCTTACCAATATCGACGGGATATCGTTGAGGACCGCGAGAAAT	6598
Db	206066	CTAGGAATGCTTTTCAGTTTAACTAACTATGCTCGGATATTTATGAGGATGCGAAGGCC	2061255
QY	6599	GGCGCTTCTATCTGSCCAAAAGCTGGCTGATGAGGCGGACTGAGCCCGCCAGCTT	6658
Db	206126	GGCGCTTCTATCTTCGCTAGAAATGCTTACCAAGAGGCTAATGCGACACACTT	2061855
QY	6659	GCCGATCCGCAACATGCGCGACGCTGGCCCCGCTGGCAGCGCGTCTGGTGGCGGAGCC	6718
Db	206186	ATTATTACCGAAACCGTCTGCGCTAGCTCGGATGCGATCCCGATTGATTGGAGAGCG	2062454
QY	6719	GAGCGTCTATCTACATGACGCGCGGACGCGGCGTCCAGAGATTGCGGCTCGGTCGCGCTGG	6778
Db	206246	GAATTTTATTATCTTCGCTTGAACAGGACTTTGGTTTACCTTTGCGTTACGCTTGG	2063055
QY	6779	GCGATGCGCACCGCGCGCGGCGCTTTTACCGGAAATTTGGCGTAAATGTCAGCATGCGCGT	6838
Db	206306	GTCATCGCTTTCGCTCATGATGATCTATCGTGAATTCGGCATTAAGTGCACAAAGCAGGA	2063655
QY	6839	GCCCGGCAATGGGATACGCGCCAGCGCACAGTAAAGCGAAGAGTGCGCGCTGCTGGTG	6898
Db	206366	GTAACGGGCTTGGGATGACAGAAAGAACTAATCGTGAGAGAAAGTTACTTATTAAAG	2064255
QY	6899	AAAGTGGCGGCGTGGCGCTTACTTTCGCGCTTGTGTCATCCGAGGCGGCTCGCGGT	6958
Db	206426	GCTGGCGGAATGAAAGCGCTTATATCTGGGGTGGGATATCAAAAGCGCTGCGATCCCAA	2064855
QY	6959	CTGTGGCAGCGCTCGCGTTGACA	6981
Db	206486	TTGTGGCAGCGACTCATTAATA	206508
RESULT 9			
ERMCRT1B	2415 bp	DNA	linear
LOCUS	2415 bp	DNA	linear
DEFINITION	Pantoea agglomerans phytoene dehydrogenase (crtI) and prephytoene pyrophosphate synthase (crtB) genes, complete cds.		
ACCESSION	M38423		
VERSION	M38423.1	GI:148401	
KEYWORDS	Pantoea agglomerans		
SOURCE	Pantoea agglomerans		
ORGANISM	Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales; Enterobacteriaceae; Pantoea.		
REFERENCE	1 (bases 1 to 2415)		
AUTHORS	Armstrong G.A., Alberti M. and Hearst J.B.		
TITLE	Conserved enzymes mediate the early reactions of carotenoid biosynthesis in nonphotosynthetic and photosynthetic prokaryotes		
JOURNAL	Proc. Natl. Acad. Sci. U.S.A. 87, 9975-9979 (1992)		

ORIGIN

NDIPERYPLQLBGRMDVEGRYDILDTLTYCYHAGVGVMMALINGARDEVLD
RASDLGAFQUTNIAARDVMDPAIGRIYLAFAFLVAGVBEAIIAAEENHAAVAGVA
RLIDVABPYDOLALIGMAALPRRAAABVAGVRYAIGIEVRKRGKANDTRVSTST-
AQAAGFLAGGLAOLALATRFKAPPFPAHLMTRPA"

Query Match 13.3%; Score 1156.4; DB 1; Length 3801;
Best Local Similarity 58.9%; Pred. No. 7e-207;
Matches 2123; Conservative 0; Mismatches 1436; Indels 45; Gaps 6;

QY 3421 GATGTCATTTCTGTGCGTGCCTGCGCTGATACGAGGCTGATGGCTGCTTACGCGAG 3480
DB 173 GACATCGTCTTCTGCGCGCGGACTTGCCATTTAGCTTATGCGGCGCGCTCGCGCC 232
QY 3481 CTGAGCGCCGCACTTAAGTTTGTCTACTGAGAGTGAAGCGGCGCGCGCGCAATCAT 3540
DB 233 CAGGTCCTCGGCTTCAATGCTGCTGTGAAGCCGCGAGAGCGTGGGCGCAATCAC 292
QY 3541 ACCGTGCTTCCATCCGAAAGAGTCAAGCGAGCGAGTTTGGTGGCTGAGCGCTG 3600
DB 293 AGCTGTCTCGCCACGACAGAGATCTCACTGCGCGAGCGCGCTTCTGCTCCCTTC 352
QY 3601 CTTTCGCGCGCTGCGCGCTGATACAGTACGCTTCCCACTGCGTCCGCACTGAGT 3660
DB 353 CAGAGCTATCTGTGCGCGCGGCAAGGAGTGAATTTCCGCTTTTTCGCGAGCTCAA 412
QY 3661 GGTGAATTTGCTGATTTGCTGCGAGAGATTTTCCCGGCACTTACAGAGTCTGCT 3720
DB 413 GGGGCGCTACGCACTATTTCTCCGAGCGATGCGCGAGTGAATGAGCGCTGTGCG 472
QY 3721 GCGCGCTACGCACTCGAGCGCGCGCTCAGAGAGTCTCACCACCGGAGGTCAAGCTGG 3780
DB 473 GCGGCTATCCGCACTATGCGCGCGCTGCGCTATGCGCGCGCGAGCGCGCTGAG 532
QY 3781 GATGCGCGAGTGTTCAGCGCGAGCGGATGATGACGAGCGCGGCTGACCGCAACCG 3840
DB 533 GGGGCGCGAGCATCAAGCGCGCGCGCTGAGAGCGGAGCGCGCGCTGCGCTGCG 592
QY 3841 CATCTGAGCTCGGCTATCAGGCACTTTGTGCTGAGAGTGGCACTGGCGCGCGCAT 3900
DB 593 CATCTGAGCTCGGCTATCAGGCACTTTGTGCTGAGAGTGGCACTGGCGCGCAT 652
QY 3901 GGCCTGAGAGCAATATGATGAGAGCGAGCGCTGATGACGAGAGGTTATCGCTT 3960
DB 653 GGCCTACCGCGCCCATCATATGATGAGAGCGAGTGAAGAGCTGGCGGTTACCGCTTC 712
QY 3961 GTTTACACCTGCGCTCAGTGGCAGCGCTTATGATGATGATGATGATGATGATGAT 4020
DB 713 GTCTATGCTGCTGCGCTGAGAGTGAAGAGCTGCTGATGATGATGATGATGATGAT 772
QY 4021 CATGCACTGATGAGTGCAGAGCGCGCGCTGATGATGATGATGATGATGATGATGAT 4080
DB 773 GGCCTGAGTCTTCCGCGAGTGCCTTGGCGCGCGCATTTCCGCTATGCGCGCGAG 832
QY 4081 GGCCTGAGTCTTGGCGCGAGTGCCTGCGAGAGAGAGCGCTGCTGCTGCTGCTGCTG 4140
DB 833 GGCCTGCGCGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 892
QY 4141 GGCATATCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4197
DB 893 GGGGAGATCAAGCGCTTTCTTGGCGAGAGCGCTTCCGCGCGCGCGCGCGCGAGT 952
QY 4198 GCGGAGCTGTTTACGCGCACCGGTTACTGCTGCGCGCGCGCGCGCGCGCGAG 4257
DB 953 GCGAGCTGTTTACGCGCACCGGTTACTGCTGCGCGCGCGCGCGCGCGAGT 1012
QY 4258 AAGATGCGAGAGCGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 4317
DB 1013 TCGGTGAGCGCGCTTCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCAT 1072
QY 4318 GCGCGTACGACTGCGCGAGAGCGCTTTTCCGCTGTTAAATGCGATGCTGTTCTT 4377
DB 1073 GCGGCGCGCGCGCGCGCGCGCGCTTCTTCCGCGCTGCTCAACCGCATGCTGTTCCG 1132

QY 4378 GCGGAGCGGCTGAACAGCGCTGCGCGCTGATGAGCGCTTTTACCGGCTTGAACCGCGA 4437
DB 1133 GCTGCGCATCGGAGCGCGCGCTATGCTATCTCAAGGCTTCTCAAGGCTGTGCGAAGAC 1192
QY 4438 TTGATTAAGCGCTTTTACGCGCGGCAACTGCGCTGAGCGATGAAGACGATCTGTGC 4497
DB 1193 CTATGCGCGCTTTCTACGCGCGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1252
QY 4498 GCGAAGCGCGGCTCTTCCGCGAAGCGCTGCGCGATGATGATGATGATGATGATGAT 4557
DB 1253 GCGCGCGCGCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1312
QY 4558 CCAAGGAGAAATATG-----AAAGCACTATGATGATGATGATGATGATGATGATGAT 4596
DB 1313 CCGGTTCCCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1372
QY 4597 CAGGCTTGGTGGCTGCGCGCTGCGATGCTGTCAGAGCGCGCGCGCTGCGCGCTGCG 4656
DB 1373 CCGGCTTGGCGGCTTCTCCCTGCGCATGCGCTGCGAGGCGCGCGCATCGCACACCG 1432
QY 4657 TGCTGAAAGCGCGATGAAGCTTGGCGGCGCGCTATGATGATGATGATGATGATGATGAT 4716
DB 1433 TGTTGAGCAGCGGAGCAAGCCGCGCGCGCTGATGATGATGATGATGATGATGATGAT 1492
QY 4717 CTTTGAATGCGGCTGCGAGCGTATGATGATGATGATGATGATGATGATGATGATGATGAT 4776
DB 1493 CTTTGAATGCGGCGCGCGCGCGCGCTGATGATGATGATGATGATGATGATGATGAT 1552
QY 4777 TGGCAGGCAAGCAACTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4836
DB 1553 CCGCGCGCGCGCAAGCTGTCCGATGATGATGATGATGATGATGATGATGATGATGATGAT 1612
QY 4837 TGTGCTGGAAGACGCGAGCGAGCTGATGATGATGATGATGATGATGATGATGATGATGAT 4896
DB 1613 TGCTGAGCGGAGCGCGCGAGCTGATGATGATGATGATGATGATGATGATGATGATGAT 1672
QY 4897 AGATGCGCACTTTTATGCGCGAGAGTGTGCGCGCTGATGATGATGATGATGATGATGAT 4956
DB 1673 AGATGCGCACTTTTATGCGCGAGAGTGTGCGCGCTGATGATGATGATGATGATGAT 1732
QY 4957 AGATGCGCACTTTTATGCGCGAGAGTGTGCGCGCTGATGATGATGATGATGATGATGAT 5016
DB 1733 AGGCGCTTTCAGAGAGGTTATCTCAAGCTGCGCGCGCTGATGATGATGATGATGAT 1792
QY 5017 ACATGCTGCGTGCAGCGCGCAAGCTGCGCTGCGCGCGCTGCGCGCGCTGCGCGCGCT 5076
DB 1793 GCAATGATGAAGCGCGCGCGCAAGCTGCGCGCGCTGCGCGCGCTGCGCGCGCTGCG 1852
QY 5077 TGTGCGCAATTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 5136
DB 1853 TGTGCTGCGCTTATGCGCGAGATGATGATGATGATGATGATGATGATGATGATGATGAT 1912
QY 5137 TGTGCGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 5196
DB 1913 TGTGCGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1972
QY 5197 GCGAATGCGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 5256
DB 1973 GCAATGAGGCGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2032
QY 5257 GGTGCTTTGCGGATTTGCGCGGATGATGATGATGATGATGATGATGATGATGATGATGAT 5316
DB 2033 AGCTGTTACCGAGCTTCCGCGGAGATGATGATGATGATGATGATGATGATGATGATGAT 2092
QY 5317 CCGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 5376
DB 2093 TGAAGGCG 2152
QY 5377 CCGTTGCGTCAATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 5436
DB 2153 TCGTCCGCTCAATGCGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 2212

QY 5437 CGGCGGTAAGCGGCGCAACGCTGAAGCGAGCGGATGAGCACTCGCTTTTGATC 5496
 Db 2213 GCGGTCTCGTAAGGCGCGGCTGGCCAGACGCGCACTCATGTGCTGTTCTCA 2272
 QY 5497 TCTATTTTGGCTTTAATCTATGCGCAACCGGAGCTGGGCAACACGATGCTTTGGTC 5556
 Db 2273 CTTATTTTCGGGCGCGGCGACCTGGGACATCTCCAGACACACACGGTGTGTTCCGTC 2332
 QY 5557 CGGCTATCTGTAATTTGATGATGATCTTCAATAGACGACAGCTGGCGGAATTTCT 5616
 Db 2233 CGGCTATCTGGGCGCTGTGAGCAAGATCTTCAAGGGCGCCACCTTCCGACACATCT 2392
 QY 5617 CGCTATCTGATGCGGCTGTCTCCAGCGATCCGCTGCTGGACCGGCGGCTGGCA 5676
 Db 2293 CGCTATCTTCCAGCGGCGCACTGGACAGCAATCGCTGGCGCGGAAGCTGTACCG 2452
 QY 5677 GTTTTACGTGCTGAGCGCGGCTGCGCATCTCGGTACCGCGCAATTGATCGGCAACAG 5736
 Db 2453 CTTTCTAGTGTCTCGCGCGGCTGCGCATCTGGGCAAGGCGGACATCGATGGAGGTGG 2512
 QY 5737 AAGGCGCGGCTTGGCGGATGCAATTTTGTCTATCTGAGAGACATATATGCGGCTC 5796
 Db 2513 AAGGCGCGCTTATCGGACCGCATCTCGCCCATCTGGAAGCGGCTACTGCGCGGCG 2572
 QY 5797 TGGACAGCAGTTATGTGACACCGTATGTTTACCGGTTGATTTTGGGACACGCGTGC 5856
 Db 2573 TGGCGACAGCTGTGACCTCGGCACTCTCAACCCCTTGGCTTCCGCAACGATCTGT 2632
 QY 5857 ACGGCAATCAGGGCTGACGCGTTTTCGCTGCAACCCATTTTACGCAAGCGCTGTTC 5916
 Db 2633 CCGGCAATCAGGGCTGCGCTTCTCGGTGAAGACCGCTCTCACCGAGTCAAGGTTC 2692
 QY 5917 GCGGCAATCAGGGCTGCGCTTCTCGGTGAAGACCGCTCTCACCGAGTCAAGGTTC 5976
 Db 2693 GCGGCAATCAGGGCTGCGCTTCTCGGTGAAGACCGCTCTCACCGAGTCAAGGTTC 2752
 QY 5977 CCGGCGCGGCTGAGCGGCTGATCGGCTCGGGAAGACCGCGGCTGATGTTG 6036
 Db 2753 CCGGCGCGGCTGAGCGGCTGATCGGCTCGGGAAGACCGCGGCTGATGTTG 2812
 QY 6037 AGGATCTGACCG-----GATGAACCAACCGCGCTGATGACGAG-----TCAC 6081
 Db 2813 CGGATCTGAGCGGCTGACCGCGGATGAGCACGCTGCGCTACCGTCTCGGTGACAGGA 2872
 QY 6082 GCAAAACATGCGGCGGCTCCAAAAGTTTGGCCAGGCTACCGGCTATTTGATCTTC 6141
 Db 2873 GGAAGCCATAAACAAGGCTCCAAAGGCTTCCGCGCGCGCTGTTCCGCGCG 2932
 QY 6142 AACCGCGCGGCTGAGCTGATGCTGTACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 6201
 Db 2933 CATGCGGGAAGACCGGCTGATGCTGTATGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2992
 QY 6202 TGGTCAAGCGCTGAGCGGCGGCGACGACGCGGCTGCGGATGACAGCGCGGAT 6261
 Db 2993 CGGCGAGGAACTGGGGGACGCGGCGGTCTCCAGCGCGCTGCGCGG---CGAGCGCT 3049
 QY 6262 GCGGCACTGCAATTCGAAACCGCGCGGCTTCAAGCGGCTCCACATGATGAACAGC 6321
 Db 3050 CGATGCGCTCTATGAGAGACCGCGAAGCGCTATCGCGCGCGCTCCGCGCATCCGCG 3109
 QY 6322 GTTTGCTGCTTTGAGAGAGGCGGCTGAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTG 6381
 Db 3110 CTTGCGCGCTTTCGCGGAGGTGAGAGCGCAACGACATTCGAGGCGCTATCCCTTCA 3169
 QY 6382 TCAATCTGGAAGGTTTGGAGTATGCGCTGGAAGACGTTATGCTGTTTGGGGAAC 6441
 Db 3170 ACTGCTGGAAGGCTTTCGCGATGAGAGTGAAGGCGCGGCTACGACACCTCGATGAC 3229
 QY 6442 GCTGCTTACTGCTATCACTGAGCGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 6501
 Db 3230 GCTCACTATGCTATCACTGAGCGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3289
 QY 6502 CGTACGTGATGAGCGCGCTACTGATCAAGCGCTGATTTGGGCTGCGGCTTACGCTTAC 6561

Db 3290 CGGCGGGAATGAGTGTGCTGAGACCGGCGGAGCGACCTCGGCTGCTCCAGCTCAC 3349
 QY 6562 CAATATGACCGGATATCTGTTAGACGCGGAGATGCGCTTGTCTATCTGCCAACAAG 6621
 Db 3350 CAATATGACCGGATATCTGTTAGACGCGGAGATGCGCTTGTCTATCTGCCAACAAG 3409
 QY 6622 CTGGCTGATGAGGCGGAGCTGAGCGCGCGCCAGCTTGGCGGATCCGCAATCGCGAGC 6681
 Db 3410 CTTCTCTGCAAGAGCGCGCTGCGGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 3469
 QY 6682 GCTGCGCGCGCTGAGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 6741
 Db 3470 GGTGCGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3529
 QY 6742 CAGCGGCTGCGGAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 6801
 Db 3530 CATCGGATGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3589
 QY 6802 TTACCGGGAATGAGCGGCTGTAAGAGTACGAGTGCAGCGGCGGCGGATGAGTACG 6861
 Db 3590 CTATCGCGCATGCGCATGAGTCTGTAAGCGGCGGCGGCGGCGGCGGCGGCGGCG 3649
 QY 6862 GCGCACCATGTAAGCGGGAAGTGCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 6921
 Db 3650 CTCACCTCAGCGGCGGCGGAGGCGGCTTCTGCGGAGTGCCTGCGGAGGCGCTCGC 3709
 QY 6922 TTGCGG---CCTTGTATATCCGAGGCGGCTGCTGCGGCTGCTGCTGCTGCTGCT 6978
 Db 3710 CACCGGCTTCAAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 3769
 QY 6979 ACAC 6982
 Db 3770 ACCC 3773

RESULT 11
 AF218415
 LOCUS
 DEFINITION
 Bradyrhizobium sp. ORS278 [geranylgeranyl synthase (crtE), lycopene cyclase (crtY), phytoene desaturase (crtI), phytoene synthase (crtB), and beta-carotene ketolase (crtW) genes, complete cds.
 AF218415
 AF218415.1 GI:8650414
 ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 Bradyrhizobium sp. ORS278
 Bradyrhizobium sp. ORS278
 Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales; Bradyrhizobiaceae; Bradyrhizobium.
 REFERENCES
 Hannibal L., Lorquin J., D'Orto J., N.A., Garcia N., Chaintreuil C., Masson-Boivin C., Dreyfus B. and Giraud E. Isolation and characterization of cantaxanthin biosynthesis genes from the photosynthetic bacterium Bradyrhizobium sp. strain ORS278 J. Bacteriol. 182 (13), 3850-3853 (2000)
 JOURNAL
 MEDLINE
 PUBMED
 20309720
 10851005
 2 (bases 1 to 6635)
 REFERENCES
 Giraud E. and Angles d'Orto J., N. Direct Submission
 Submitted (02-SEP-1999) Laboratoire des Symbioses Tropicales et Méditerranéennes, TA/10J, Campus de Baillarguet, Montpellier 34398, France

FEATURES
 source
 location/Qualifiers
 1. 6635
 /organism="Bradyrhizobium sp. ORS278"
 /mol_type="genomic DNA"
 /strain="ORS278"
 /db_xref="taxon:114615"
 complement (273..1223)
 /gene="crtE"
 complement (273..1223)
 /gene="crtE"

```
/codon_start=1
/transl_table=11
/product="geranylgeranyl synthase"
/protein_id="AA078199.1"
/db_xref="GI:8650415"
/transl_table="MKEKVDLTPDAFETOLDKRRRIGSEAVANAMFGTTTAPICQ
GSHAVLAGKRRKGMVYALGSDLGYPEQGLSSAVITTHASIVYDLDKMDA
RRRSOPATVAVGEATATLSSILIRAEVVARQDSFASRSSIVDTLSHAIGQ
ALCGQYDLYPPYVATEODLIHRVORKEALFYAARCPALLAEVDPELILIRAG
ORLGVAFQIPDDLLDTGDMAHIGKVDGQDHGTVATLLGLAPARAERADDEIAVQK
ELRETVGGRALDIRMAARIAGTGKKSAGRDLRPHAG"
1446..2630
/gene="crtY"
1446..2630
/gene="crtY"
/codon_start=1
/transl_table=11
/product="lycopen cyclase"
/protein_id="AA078200.1"
/db_xref="GI:8650416"
/transl_table="MSRDADVIYVGGAGLCLALRLTARPDRLRVIIEGSASIAQ
HTWSFPGTIDISSQHWLGRVHWPGEYERFAEHAIRLSTAYLSMTSTRLEAVQ
RPERILRDATASATADHYVEGGRTLRACVLDAGGRVPRGLAGFOKGLAEVR
LYAPHGIDVPIVMDATVAOSDGYRFVTLPLDPORLLIEDTVYSDGDELPEOYLHOB
ARALAKGMQIAEIRAEQSVLPYIAGDSGLVSKRSPRNGALMLPHTTGYSL
PDARVADLLPARIQKGLSSADARETIDGKRTWRKGYRFLRMLPKAEBS
RHRILARFGLDQALIRFYAARIQPDKLRFEMMLMKPPISSALACLPEASAK
TP"
2627..4147
/gene="crtI"
2627..4147
/gene="crtI"
/codon_start=1
/transl_table=11
/product="phytoene desaturase"
/protein_id="AA078201.1"
/db_xref="GI:8650417"
/transl_table="WSQEMQAKTAVVIGSGFGLSLAIRLQSGAIATTLVERKDFG
GRAVYQDGFEDPAGPTVITDPTCLQELFALSRLKLENYELMPVPFYOLRMDA
TPYVNDQALERQIAAFCPADVDGYRFRFSYERLEEGVYKLGHPFDPFSMRYV
APOLVALQSYRSYKSOYVSDHLEQAQSFSLVGNPFAISSYALILHLERW
GWPPRGGTALINKGLVOLPKDIGEVTLSYSRIETANGRVSAVADGRPAD
IVASNDVYVTRDLDKDBELPRTASLKRKRSMSLPYIRGLRHEHLEKHITL
FGRVYRELINIEFKGPALPEPFSYLHAPSVTDSLAPQGCSTYIVSLVPVHAAAT
DWSVEGRYRDRILDYLEARILPGLKSDLATCRIFTPQDNTENLNAHLSAFSLERPL
TQSAVFRANHADDKIKGLVYVAGTHPGAGIPGVGSAKAKARVILDEQBLEAQ"
4144..5151
/gene="crtB"
4144..5151
/gene="crtB"
/codon_start=1
/transl_table=11
/product="phytoene synthase"
/protein_id="AA078202.1"
/db_xref="GI:8650418"
/transl_table="WLRNDNDHLARLSEAVIRNGSKSFAAASKLPDSRTASVHLTYA
WCRHCDVIDGQDGIROGVGAPGPQIGTQMLRDQALBEGAPMDPVFOGLORV
OEHAIPHHVFEILDGPAMDVGREYETLSETLDYCYHAGVGVVMSALMGAREEAT
LDRPADGIALQLTNIARVDIEDQGRMYLPQOMLCEAGVPAEVAEVRORVAFV
VARLDVAEOFEASEAGIARLPVRCMAVETARVYVROIGRVEMKRGPGAMARLAT
TGAOKLGAIGRSALTGLTRTWVYVPAERHNLWTRPARRAASDAQASFANACSNRG
Y"
5123..5899
/gene="crtW"
5123..5899
/gene="crtW"
/codon_start=1
/transl_table=11
/product="beta-carotene ketolase"
/protein_id="AA078203.1"
/db_xref="GI:8650419"
/transl_table="MHAATAKATFRGASRRDDARORRVGLTAAVITIAAVLVHVGIM
FPMPLTSHSLPLPLVYVLOTWLYVGLPIIAHDMGSSLPFRKQVRRIRIGQLCTFLY
```

ORIGIN		Query Match		12.5% Score 1095.8; DB 1; Length 6635;		Best Local Similarity 58.2% Pred: No. 1.3e-193;		Matches 2075; Conservative 0; Mismatches 1427; Indels 63; Gaps 7;	
AGSPDALNVEHHKHRRHCTAEDPDDEVPHPFMWMAFPFLHYGKQVATAAV SLVYQLVFAVPLQNIILLFMALPGLSLALQILFTTGTYLPKRPATQPPADRINATSEEP AMLSLLTCHFGHFBHHHLIPDAPWMLPEIKRRALERD"		Query Match		12.5% Score 1095.8; DB 1; Length 6635;		Best Local Similarity 58.2% Pred: No. 1.3e-193;		Matches 2075; Conservative 0; Mismatches 1427; Indels 63; Gaps 7;	
Qy	3386	GTGCTGGCGGAGGCACTGATCGCGGCACCTTATGATGTCATTTGCTGTCGCCGCC	3445	3386	GTGCTGGCGGAGGCACTGATCGCGGCACCTTATGATGTCATTTGCTGTCGCCGCC	3445			
Db	1426	GAGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1485	1426	GAGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1485			
Qy	3446	TGGCTAAAGGCGGCTGATGTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	3505	3446	TGGCTAAAGGCGGCTGATGTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	3505			
Db	1486	TTGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1545	1486	TTGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1545			
Qy	3506	TACTGAGAGTCAAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG	3565	3506	TACTGAGAGTCAAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG	3565			
Db	1546	TCAATCGAGGCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG	1605	1546	TCAATCGAGGCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG	1605			
Qy	3566	TCAAGCAAGCGAGTTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	3625	3566	TCAAGCAAGCGAGTTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	3625			
Db	1606	TCTGTCGCAACAGACAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1665	1606	TCTGTCGCAACAGACAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1665			
Qy	3626	AGGTACGCTTCCCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	3685	3626	AGGTACGCTTCCCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	3685			
Db	1666	AGGTGCGCTTCCCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1725	1666	AGGTGCGCTTCCCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1725			
Qy	3686	AGGATTTGCGCGGCACTTACAGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	3745	3686	AGGATTTGCGCGGCACTTACAGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	3745			
Db	1726	CGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1785	1726	CGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1785			
Qy	3746	TCAAGCAAGTCTGACCCAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG	3805	3746	TCAAGCAAGTCTGACCCAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG	3805			
Db	1786	CCATCTCGGCAAGGCGGCAACATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1845	1786	CCATCTCGGCAAGGCGGCAACATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1845			
Qy	3806	CGGTGATTCAGCGAGCGGCGGCTGAGCGGCAACCGCATCTGCAAGCTTGCTATG	3865	3806	CGGTGATTCAGCGAGCGGCGGCTGAGCGGCAACCGCATCTGCAAGCTTGCTATG	3865			
Db	1846	CGGTGATTCAGCGAGCGGCGGCTGAGCGGCAACCGCATCTGCAAGCTTGCTATG	1905	1846	CGGTGATTCAGCGAGCGGCGGCTGAGCGGCAACCGCATCTGCAAGCTTGCTATG	1905			
Qy	3866	TTGTCGCTGAGGAGTGGCACTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG	3925	3866	TTGTCGCTGAGGAGTGGCACTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG	3925			
Db	1906	TCTCGGCTGCGAGGCTGCGGCTGCGGCGCGCGCGCGCGCGCGCGCGCGCGCG	1965	1906	TCTCGGCTGCGAGGCTGCGGCTGCGGCGCGCGCGCGCGCGCGCGCGCGCGCG	1965			
Qy	3926	ACCGCAAGCTGATCAGCAGCAGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTG	3985	3926	ACCGCAAGCTGATCAGCAGCAGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTG	3985			
Db	1966	ACCGCAAGCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG	2025	1966	ACCGCAAGCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG	2025			
Qy	3986	GCGCTTACTGATTAAGATACCACTACATCAACATGCGCAGCTGCGATGCGCACA	4045	3986	GCGCTTACTGATTAAGATACCACTACATCAACATGCGCAGCTGCGATGCGCACA	4045			
Db	2026	AGCGGCTGTTATTCAGAGCACTACTATAGCAGCGCGCGCGCGCGCGCGCGCG	2085	2026	AGCGGCTGTTATTCAGAGCACTACTATAGCAGCGCGCGCGCGCGCGCGCGCG	2085			
Qy	4046	CGGCGGCTGCACTTACGAGATTATGCCACAGCGCGCGCGCGCGCGCGCGCGCG	4105	4046	CGGCGGCTGCACTTACGAGATTATGCCACAGCGCGCGCGCGCGCGCGCGCGCG	4105			
Db	2086	TGCAATCAGCGCATCGCGGCTGCTGCGGCTGCCCAAGGCTGCGAGATGCCAGAT	2145	2086	TGCAATCAGCGCATCGCGGCTGCTGCGGCTGCCCAAGGCTGCGAGATGCCAGAT	2145			
Qy	4106	GCGAGAGCAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	4165	4106	GCGAGAGCAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	4165			
Db	2146	GCGAGAGCAGGAGTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	2205	2146	GCGAGAGCAGGAGTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	2205			
Qy	4166	AGCAGCAGGCGCAACGCTGCAAGCGGCTGCGCGCGCGCGCGCGCGCGCGCG	4225	4166	AGCAGCAGGCGCAACGCTGCAAGCGGCTGCGCGCGCGCGCGCGCGCGCGCG	4225			
Db	2206	AGCGCGCAGCG	2265	2206	AGCGCGCAGCG	2265			
Qy	4226	ACTGCGTCCG	4285	4226	ACTGCGTCCG	4285			
Db	2266	ATTGCTGCGCGAGAGCGGCTGCGGCTGCGGATCTGCTGCTGCTGCTGCTGCTG	2325	2266	ATTGCTGCGCGAGAGCGGCTGCGGCTGCGGATCTGCTGCTGCTGCTGCTGCTG	2325			
Qy	4279	GCGGACGCTGACAGCTGA-----GCCATGCAATGCAATCTTTGCCCGTCAGCA	4333	4279	GCGGACGCTGACAGCTGA-----GCCATGCAATGCAATCTTTGCCCGTCAGCA	4333			
Db	2326	GCGGCGCTCTGAGAGGCTGAGCGCGGAGAGCATGCTAGCTAGCGCGGAGCAT	2385	2326	GCGGCGCTCTGAGAGGCTGAGCGCGGAGAGCATGCTAGCTAGCGCGGAGCAT	2385			

Db	LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS
Db	4546	CTGAGACTACTCTATCACTGAGCCGGCGGTGGTCGGCGCGATGATGATGCGGCATCATGCGGC	4605		
Qy	6503	GTACGTGATGAGCCGGCTACTCGATCAACCCCTGTGATTTGGCTTGCGGCTTTCAGCTTAAC	6562		
Db	4606	GCTCGCGAGAGGGCGACCGCTGGACCGCGCGCGCCGATCTCGGCATTCGCGTCCAGTCCAC	4665		
Qy	6563	AATATCGCAGCGGATATGCTTGAAGGAGCGCGGAGAAATGGCGCGTTCATCTGTCGCAAAAGC	6622		
Db	4666	AACATCGCCCTGACGTATGAGGATATCCAGACCGGCGCCCATGTATCTGCGGAGCA	4725		
Qy	6623	TGGCTGATGAGCGCGCATGAGCGCCCGCAGCTTGGCCGATTCGCAACATCGCGCAGCG	6682		
Db	4726	TGGCTGTGCGAGGCGCGCGCTCGCGCGCCCGCAGATCGCGGAAACCGCAGCATTCGCGAGCG	4785		
Qy	6683	CTGCGCCCGCTGGCAGCGCGTCTGTGTGCGGAGACCGGATCTATCACTAGCGCGC	6742		
Db	4786	GTGTTTCCGTGTCGTGCGCGCGCTGCTCGATGTGCGGAGCAGTTTACGAGGCGAGCG	4845		
Qy	6743	AGCGGGCTGCGCAGATTTGCGCGCTCCGTTCCGCGGTGGGATTCGCAACCGCGCGGGCGTT	6802		
Db	4846	GCAGGCAATCGCCCGCTGCGCGGTGTTGCGCGCTGGGCGGTGAGAACCGCGCGGTGCTC	4905		
Qy	6803	TACCGGGAATTTGGCGTAAAGTGACAGATCGCGGTGCGCGGATGAGGATACGCGCAG	6862		
Db	4906	TATGCGCAGATCGGCGCGGAGATGATGAAGCGCGGTCCGCGCGCTGGAACCGCGCATC	4965		
Qy	6863	CGCACCACTAAAGCGGAAAGCTGG	6887		
Db	4966	GCCACGACAGCGCGCGCAGACCTGG	4990		
RESULT 12	PMCRWTZYI	6335 bp DNA linear BCT 15-SEP-1999			
LOCUS	Paracoccus marcusii crtW, crtZ, crtY, crtI, crtB & crtE genes.				
DEFINITION	Y15112				
ACCESSION	Y15112.1	GI:5912291			
KEYWORDS	astaxanthin biosynthesis; beta-carotene C-4 oxygenase; carotene hydroxylase; crtB gene; crtE gene; crtI gene; crtW gene; crtZ gene; geranylgeranyl pyrophosphate synthase; ketolase; lycopene beta-cyclase; phytoene desaturase; phytoene synthase. Paracoccus marcusii				
SOURCE	Paracoccus marcusii				
ORGANISM	Bacteria; Proteobacteria; Alphaproteobacteria; Rhodobacterales; Rhodobacteraceae; Paracoccus.				
REFERENCE	1	Harker, M. and Hirschberg, J.			
AUTHORS	Carotenoid biosynthesis genes in the bacterium Paracoccus marcusii M1				
TITLE	Unpublished				
JOURNAL	2 (bases 1 to 6335)				
REFERENCE	Hirschberg, J.				
AUTHORS	Direct Submission				
TITLE	Submitted (12-SEP-1997) J. Hirschberg, The Hebrew Univ of Jerusalem, Dept of Genetics, Life Sciences Institute, Jerusalem 91904, ISRAEL				
JOURNAL					
FEATURES	location/Qualifiers				
source	1. 6335				
	/organism="Paracoccus marcusii"				
	/mol_type="genomic DNA"				
	/strain="M1"				
	/db_xref="taxon:59779"				
	409..1137				
gene	/gene="crtW"				
	409..1137				
	/gene="crtW"				
CDS	/function="astaxanthin biosynthesis"				
	/codon_start=1				
	/transl_start=1				
	/product="beta-carotene C-4 oxygenase (ketolase)"				
	/protein_id="CAB5605.1"				
	/db_xref="GI:5912292"				
	/db_xref="GOA:Q9RLH7"				

gene /db_xref="rEMBL:O9RLH7"
/translation="MSAALPKALDTLSTLVSGLTAAVLHVALHMLFLDAANPIL
LANAPFLWTLSVGLFPIANDAHGSGVPRPANAAMQGLVIMLYAGSPKRCIKYK
HNAHNRHAGTDDEDDPDHGGFVWRVARTIGYFWMRGGLLPVIVTAYALILGRMMY
VFWMLPGLILASIQLFVPGTWLPHRPGHDAPDRHNRASRIIDPVSILTCFHGGVYH
HEHHLPVPMWRRLPSTRTKGDTA"
1134. 1622
/gene="crt2"
1134. 1622
/gene="crt2"
/function="astaxanthin biosynthesis"
/codon_start=1
/transl_table=11
/product="carotene hydroxylase"
/protein_id="CAB56060.1"
/db_xref="GI:5912293"
/db_xref="GOA:O9RLH6"
/db_xref="rEMBL:O9RLH6"
/translation="MTNPLIVATVLYVWELTAVSVHRIMHGLGCMGHKSHHEHNDH
ALEKNDLGLVFAVATVTLFTVGIMAPVLMVIMLGMTYGLIFVLDGVYHQRNPP
RIPIPRKGARRLYQAHRLHAAVEGRDCHVSFGLYAPVDKLKDLTKTSGLVRAEQE
RT"
1619. 2779
/gene="crtY"
1619. 2779
/gene="crtY"
/function="astaxanthin biosynthesis"
/codon_start=1
/transl_table=11
/product="lycopenene beta-cyclase"
/protein_id="CAB56061.1"
/db_xref="GI:5912294"
/db_xref="rEMBL:O9RLH5"
/translation="WTHDVLLAGAGLNGLIATLALRAAPDLRYTLIDHNAQSEGGHT
WSEHPDLSPIMLARKPLBRANMPDOEVRRPARRLATGSLDGAALADAVRSG
AETRPDLSLIDOGATLSCGTRIEAGSVLDGRGPHSPRLTLGFHKPLGVEIETPR
PHGVRPPIYMGTTVQDQGYCFYILPPSPRILIEPTRISDGDLDLDAALASQVY
AKQQWTAQVAREKGIPLIALAHDAQFMADHAEQVPGVLRAGFFHPVATGSLPYA
AQVAVVAGLGGPPTDALRGAIRDYAIDRARDRFLRLNRLMFCACDPRRTYLLQ
RLYRPHGLIERFYAGRLSVADQLRIYVKGPIPIGLTAIRCLPERPILKENA"
2776. 4281
/gene="crtI"
2776. 4281
/gene="crtI"
/function="astaxanthin biosynthesis"
/codon_start=1
/transl_table=11
/product="phytoene desaturase"
/protein_id="CAB56062.1"
/db_xref="GI:5912295"
/db_xref="GOA:O9RLH4"
/db_xref="rEMBL:O9RLH4"
/translation="MNAHSPAKTAIVYAGFGGLAALRIQSGIATTLVLEARDKPG
GRAYWMOQHFDAGPVTITDPDLAKMALITQDMARDVTLLMPVSPFTHLMPGSK
VFDVYDAVDQERQIAQENPDLEGYRFRFYAEVYQEGVKKGVPLKLGQMLKA
APALKLNAQSVHAKVATFTKIDYLRAPSEYHLLVLVGQPFSTSSLYALIHAEKRG
GWPFKQGTNOLVAGMVALFERPGQMMLNANKVARIETEGARTGVTLADGRSIRADMP
VASNDWHNTRVDLIGHARGOSRAKSLIDRWGMSLVLFHFLGRBAKDIANTHTILE
GRYKELYVNEIFKQKYLADPSLYXHSCTTDDPMAPGMSHTYVALPVPLISAEIIT
WASVRPRADRKTIAFLERLIPNLKANITRITRIPADPFASELAAHHSQAFSVPILIT
QASWPRPNRDKTIRNFYLVGAGTHPGAGIPGVVGSAAKTAQVMLSDLASA"
4278. 5192
/gene="crtB"
4278. 5192
/gene="crtB"
/function="astaxanthin biosynthesis"
/codon_start=1
/transl_table=11
/product="phytoene synthase"
/protein_id="CAB56063.1"
/db_xref="GI:5912296"
/db_xref="GOA:O9RLH3"
/db_xref="rEMBL:O9RLH3"

/translation="MSDLVLTSTREATITOGSGSFATAKIMPRGTRDDTWMYXACRHA
DDVIDGAGSRRPRAVNDPQARLDGADTADTAIOGCPPTPPALRAVARRRDFQ
AWEMDLIEGPAIDVARDYRTLDVLESTHAGILGVAMARVGVGDHPLDCAOL
GLAOLTHIADVIDIARIGRCPYDLMBDAGARVGPSPPLVYVILRLDAAEP
YSPARVGLADLPRCAWSIPARIVYALILPIRKGGPEAYRORISTSKAKIGLLG
IGGPDVADHACRGSVGRDMLTRPFFHNA"
complement (5189..6064)
/gene="crfb"
/complement (5189..6064)
/function="astaxanthin biosynthesis"
/gene="crfb"
/function="astaxanthin biosynthesis"
/codon_start=1
/transl_table=11
/product="geranylgeranyl pyrophosphate synthase"
/protein_id="CAB56064.1"
/db_xref="gi:5912297"
/db_xref="goa:Q9RLH2"
/db_xref="trEMBL:Q9RLH2"
/translation="MRBDVNPVIAHTLIQTRIEIAOGRGAVLOPIGAMALSGKRF
RGMILMLAAESGVCITVDACAVENKHAASLITFDLPENDAGLRROPATVHAH
GESRAVIGGILITEAMALLAGARAGETVAQOLVRLSRLOQGLCAQDLDLHAH
KMGAVGBOBODLKTGVLFIAGLEMLAVIKEPDABEQMTDFGQLRVPFQSYDLE
VVGQOALGKDTGRDAAAPGRRIGLAVSDIQKVSQRYEASRAQLIEGVSGKRLQAE
IEALLERVLFPYARA"

Query Match 12.2%; Score 1054.6; DB 1; Length 6335;
Best Local Similarity 57.1%; Pred. No. 9.4e-188;
Matches 2056; Conservative 1; Mismatches 1500; Indels 42; Gaps 6;

ORIGIN

QY 3334 CTGTAAGCGGCGCGCTGATATTGTCAGACAGCGCGTGTGTCAGACAGCAATCGTGCTGCG 3393
Db 1541 CTATGCGCGCGCGCTGACAGCTGACAGCAAGCACTGAAGACGTGCGCGCTGACGCGGC 1600
QY 3394 GAAGCGCACTGATGCGCGCACTTATGATGTCATTCTGTCGTCGCGCGCTGCTGCTAAC 3453
Db 1601 CGAAGCGGAGAGAGCGACGTAACCATGACGTGCTGTCGACAGGAGCGCGCTTGCGAAC 1660
QY 3454 GGGCTGATGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3513
Db 1661 GGGCTGATGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1720
QY 3514 AGTCAGGCGGCGCGCGCGGCAATCATCTGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCG 3573
Db 1721 CATGCGCGGCGGCGCTGCAAGGCGCATACCTGTCATGCTGCAACCGCGATCTGCTGCGC 1780
QY 3574 GCGCAGTTTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3633
Db 1781 CACTGCGTGGCGCGCTGACCCCTGCGCGCGCAACTGCGCGCGCAAGAGGTGCGC 1840
QY 3634 TTCCCGCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3693
Db 1841 TTTCCTCCGCAATGCGCGCGCGCTGCGCAACCGCTTACGCGGTGCTGAGACGCGGCGCGCTG 1900
QY 3694 GCCCGGCACTTACAGCAGAGTCTGCGTGGCGCGCTGACGACGCGCGCGCTGACGAG 3753
Db 1901 GC---GAATGCGGTGCTGCGTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTG 1957
QY 3754 GTCTACCCACCGGCGGTGACAGTGGCGAGTGGCGGAGTGTACAGCGCGACGCGGTGATT 3813
Db 1958 CTGAGTGGCGAGGGGGCGACACTGTCGCGGCAACCGGATCGAGGCGCGGTGCGTCTG 2017
QY 3814 GACGGAACGCGGCTGCAACCGCATCTGACGCTGCGCTATCAGGCAATTTGCTGCT 3873
Db 2018 GACGCGCGGCGCGCGCAACCGCTGCGCGCATCTGAGCTTGGGTTTCCAAATTTCTTGGGT 2077
QY 3874 CAGAGGCGCACTGCGCGCGCGCGCTGACGAGCAATATTGATGAGCGCGC 3933
Db 2078 GTCAAGATCGAGACCGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCA 2137
QY 3934 GTGATACGACAGCAGGTTATCGCTTTGTTTACACCTGCGCGCTACGTCAGCGCGTTTA 3993
Db 2138 GTCAACCGAGCGGAGCGGTTATCGCTTATATGCTGCTGCGCTTCTCTCGACCGCGCATC 2197

QY 3994 CTGATTGAAGATACCCCATCTATCAACCATGCCACGCTGGATGCCGACAGCGCGCGCT 4053
Db 2198 CTGATTCAGAGACAGCGCTATTCGATGAGCGCGCATGTGACAGCAGCGCTGGCGCGC 2257
QY 4054 CACATTCAGGATTAATGCGCCACAGCGCGCGCTGGAATTTGGCGCAGCTGCTGCGCGAG 4113
Db 2258 GCGTCCCAAGACTATGCG 2317
QY 4114 CACGCTGCGTCCGATCAGCGTCAAGCGCGCATATCATGATTTGCGCAACAGCAGCAC 4173
Db 2318 GGCATCTGCTATTTGCTTGCGCGCATGACGCGCGCGCGCTTCTGCGCGCATACGCGAG 2377
QY 4174 GGGCAACCGTGCAGCG 4233
Db 2378 GGGGCTGTTCCGCGGACCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 2437
QY 4234 CCGCGCGCGGTGGCGCTGCGCGGAGAAATTGCGAGC---ACGCTGCCGCGCGAGCTCAC 4290
Db 2438 CCTATGCGCGCGAGTGGCG 2497
QY 4291 ACGCTAGCCACTGATCGAATCCCTTTCGCCGTCAAGACTGGCGCGAGCAGCGCTTTTC 4350
Db 2498 GCGCTGCGCGCGCGCGCATCCGCGATTACGCGATCGACCGGCGACGCGCGCGCGCTTCTG 2557
QY 4351 CGTCTGTAATGCGCATGCTGCTTCTTCCGCGAGCGCGCTGAACAGCGCTGCGCGTATG 4410
Db 2558 CGCTGCTGAACCGGATGCTGCTTCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTG 2617
QY 4411 CAGCGTTTATACCGGCTTGAAGCGCGGATTGATGACCGCTTTTACGCGCGGCAACTGCGC 4470
Db 2618 CAGCGGTTATACCGCATGCGCGCATGATGATGAAAGGTTTATGCGCGCGCGCTGAGC 2677
QY 4471 CTAGCGATTAAGCAGCATCTGTCGCGCAACCGCGGTCTCTTCTGCGGAAAGCGCTG 4530
Db 2678 GTGCGGATCAGCTGCGCATGTCGCGGCAAGCTTCCATTCCTCTTGGCAAGCGCAT 2737
QY 4531 CGCGCATTAATGATGACCTGCGCTTACAGGGAAGAAATA----- 4572
Db 2738 CGCTGCTGCGCGGAGCGTCCCTGCTGTAAGAAACGATGAAAGCGCGCATTCGCCCGCG 2797
QY 4573 TGAAGCGCATTAATGATGAGCGCGAGCTTGGTGGCTGCGCGCTGCGCGATTCGTCG 4632
Db 2798 CCAAAACCGCATTCGATGAGCGCGAGCTTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCTG 2857
QY 4633 AAGCGCGCGCGCGCGCGTACGCTGCTGAACAGCGCATTAAGCTGCGCGCGCGCGCT 4692
Db 2858 AGTCGCGCGCGCATGCGACCAACCTGCTGAGCGCGCGCGCGCGCGCGCGCGCGCGCT 2917
QY 4693 ATGTGATCAAGATCAGGCTTTTACCTTTGATGCGCGTCCGACGCGGTGATTACGATCCCA 4752
Db 2918 ATGTCTGACAGATCAGGCGCGCATGCTTTCAGCGCGCGCGCGCGCGCGCGCGCGCGCTG 2977
QY 4753 GCGCTATCAGGCGCGCTTTTACGCTGCGAGCAAGCAACTCAGTATTAATGTCGACTGA 4812
Db 2978 ATGCGCTCAAGAGCTGTGGCGCTGACCGGGGACAGCATGGCGCGGAGCTGACGCTGA 3037
QY 4813 TGCGGTGAGCGCATTTTATGCTGCTGCTGCGGAAGACGCGAGCGAGCTGACTGACACA 4872
Db 3038 TGCGGTGAGCGCATTTTATGCTGCTGCTGCTGCGCGCGCGCGCGCGCGCGCGCGCGCTG 3097
QY 4873 ACAATCAGGCGCGCGCTGAGAGAGATGCACTTTTATTCCTCAGAGATGTCGCGGTT 4932
Db 3098 ACGAGCGCGATCACTGAGAGAGCGCAATTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCT 3157
QY 4933 ACCGCGAGTTTCTGCGCTATTCACAGATGTTTCTGTAAGGCTATCTGAACCTGGGCA 4992
Db 3158 ACGGTGCTTCTCGATTTAAGCTGAGAGGTATACAGAGGGCTAAGTCAAGCTGGGCA 3217
QY 4993 CCGTACTTTTCTGCAATTTCCGAGCAATGCTGCGTGGCGCGCGCGCGCGCGCGCGCGCTG 5052
Db 3218 CCGTGCCTTCTCAAGCTGAGCGCAATGCTCAAGAGCGCGCGCGCGCGCGCGCGCGCTG 3277

QY 5053 AGGCTGCGCAGTGTCTACAGATGGTGGGAAATTTATTCATGACGATCATCTGGCC 5112
 DB 3278 AGGCTTACATTCGCTTCATGCAAGGTGGGACCTTCATCAAGACCCCTATCTGGGC 3337
 QY 5113 AGGCTTTTCTTCTTCTACTGCTGTGTCGCGGTAATCTTTTTCACGCTTTCATCT 5172
 DB 3338 AGGCTTTTCTTCTTCTACTGCTGTGTCGCGGTAATCTTTTTCACGCTTTCATCT 3397
 QY 5173 ATACCTTAATTCACGACTGAGAGCGCAATGGGCGCTGTGTTTCGCGCGCGGTACCG 5232
 DB 3398 ATGCGCTGATCCACGCGCTGAGAGCGCGCGCGGATCTGTGTCGCAAGCGCGCA 3457
 QY 5233 GTGCGCTGTGATGAGATGGCGCGCTGTGTCGATTTGGGCGGTAATCTGTGCTCA 5292
 DB 3458 ACCAGCTGTGCGGCGGATGTCGCGCTGTTCGACGCGCTGGCGGCAAGTGTGTA 3517
 QY 5293 ACCGCGAAGTCAGCAGCTGAGACCGAGGTAACCCGATCAGCGGTGTGCTGAGAG 5352
 DB 3518 ACCGCGAAGTCGCGCGGATGAGACCGAGGCGCGGACCGAGCGGTGACCTGGCCG 3577
 QY 5353 ATGGGCGCGTTTTCGCGCGCGCGCTGTGTCATTAATCTGACGTGTGATCACTACG 5412
 DB 3578 ACCGCGCGCTTGTGAGGCGGACATGCTGCGCAGCAACGCGGACGTATGACACTATC 3637
 QY 5413 ATGCGCTGTAAAGCAGCATCTGCGCGCGGTAAAGCGCGGCAACGCTGAAGCGCAAG 5472
 DB 3638 GCGACCTGTGTGGCACACGCGCGCGGAGAGCGCGGAAATCTGTGACCGCAAGC 3697
 QY 5473 GATGAGCAACTGCTGTGTTTGTACTTATTTTGTGTTTAAATCAATGCGCAACCGCAGCTG 5532
 DB 3698 GCGGTGCTATGTGTGTTGTGTGCTGATTTGCGTCTGCGCGAGCGCCCAAGGACATCG 3757
 QY 5533 CGCACACACAGTGTGCTTGTGTCGCGCTATGCTGAATGATGATGATCTTCAATA 5592
 DB 3758 CGCATCACACATCTGTGTGCGCGCCCGCTAACAGAGCTGTCAAGATATTTCAAG 3817
 QY 5593 GCGACGAGCTGGGGAAGATTTCTGCTATCTGATCTGAGCTGTGCTGCAAGATCTCT 5652
 DB 3818 GCGCGAAGCTGGCGGAGATTTCTGCTATCTGATCTGAGCTGTGCTGCAAGATCTCT 3877
 QY 5653 CGCTGCGACCGGCGGCTGCGGAGTTTAACTGCTGCTGCGCGGCTGCGCATCTCGGTA 5712
 DB 3878 ACATGGGCGCTCGCGGATGTCACGCAATTAAGTGTGCTGCGCGGCTGCGCATCTCGGTA 3937
 QY 5713 CGCGCGCAATTTGCTGCAACAGAGAGCGCGCTGCGCGGCTGCGCATCTCGGTA 5772
 DB 3938 GCGCGCAATTTGCTGCAACAGAGAGCGCGCTGCGCGGCTGCGCATCTCGGTA 3997
 QY 5773 TGGAGGAGCACTATGAGCGGCTGCGGAGTGAAGAGTGAAGACACCGTATGTTACG 5832
 DB 3998 TGGAGGAGCACTATGAGCGGCTGCGGAGTGAAGAGTGAAGACACCGTATGTTACG 4057
 QY 5833 CGTTTATTTTTCGCGACACGCTGACGCGCATGAGGCTGAGGCTTCTCGTCAACCA 5892
 DB 4058 CGCGCGATTTTCGCGACGAGTGAAGTGAAGCGCATGAGGCGTCTTTCGCTGAGCGCA 4117
 QY 5893 TTTTGAACGCAAGCGCTGTGCTGCGCGGATTAACGCGATGCGCATTTACTTAACCTTT 5952
 DB 4118 TCTTGAACGCAATCGCGTGTGCTGCGCGGATTAACGCGATGCGCATTTACTTAACCTTT 4177
 QY 5953 ATCTGAGGCGGCTGAGCGCATTCGCGGCTGAGGCTGAGGCGTGTGCTGCGCGCA 6012
 DB 4178 ATCTGAGGCGGCTGAGCGCATTCGCGGCTGAGGCTGAGGCGTGTGCTGCGCGCA 4237
 QY 6013 AACGCGACCGCGCTGAGTGTGAGGATGTAACCGGATGAACCAACCGCGCTGATTTGA 6072
 DB 4238 AGGCGACGCGCGCGAGTGTGAGGATGTAACCGGATGAACCAACCGCGCTGATTTGA 4294
 QY 6073 GCGAGTGAACGCAACCAATGCGGAGGCTTCAAAAGTTTTCGCAAGCGCTTATTT 6132
 DB 4295 GACCTGAGCGAGGCGGATGCAACCAAGGCTTTCGCAAGCGCGCTTATTT 4354
 QY 6133 TGATCTTCAAGCGCGCGAGTACGCTGATGCTGTACGCTGTGCTGTACTGTGACGA 6192

DB 4355 GCGCGCGGATTCGCGAGCAGCAGCTGATGCTCTATGCTGTGCTGCCACCGGATGA 4414
 QY 6193 TGTGATGATGATCAGACGCTGCGGAGAGCGGACAGCAGCAGCGGTGAGATGACA 6252
 DB 4415 CGTATGACGCTCAGGCGCTG-----GGCAGCGCGCGGAGCGGATGAGACCGCA 4468
 QY 6253 GCGCGGATTCGCGAGCTGTAATTCGAAACCGCGCGCTTACAGCGGTGCTCAATGGA 6312
 DB 4469 GCGCGCGTGAAGCGCTGTGTCGACACGCTGCGCGCTTCAAGGCGAGCTGCGGT 4528
 QY 6313 TGAACGAGCTTTCGTCCTTTCAGGAAAGTGGCGCTGACGATCAGCTTCCGAGGCT 6372
 DB 4529 GAGCGCGCTTTCGCGCGCTGTCGCGCGTGGCGGAGGATGATTTCCCGAGGCTG 4588
 QY 6373 GAGCTTTCATCTGAGAGGTTTTCGATGATGAGCGCGTGAAGAACTTATGCTGTT 6432
 DB 4589 GCGCATGAGCTGATGAGAGCTTTCGATGATGATGATGAGCGCGGACTATGACGCT 4648
 QY 6433 CGGAGACGCTGCTTACTGCTATCACTGCGCGCGGTGTGAGGCTTATGATGCGCG 6492
 DB 4649 GATGACGCTGGAATATTCCTATCACTGCGCGGATCTGCGGTGATGATGCGCG 4708
 QY 6493 CGTATGAGGCTGATGAGCGGATGAGCGGATCACTGACGCTGATTTGGGTGCGGCT 6552
 DB 4709 CGTATGAGGCTGCGGACCATCTGCTGACCGCGCTGACCTGAGGCTGCGGCT 4768
 QY 6553 TCACTTACCAATATGACAGCGGATATCTTGAAGACGCGGAGATGCGCTTACTCT 6612
 DB 4769 CCACTGACCAATGCGCGCGGACGATGATGATGATGATGATGATGATGATGATGAT 4828
 QY 6613 GCGCAAGCTGCTGATGAGGCGGACCTGAGCGCGCGCTGCGCTGCGATCGCAACA 6672
 DB 4829 GCGCGGAGATGCTGATGACCAACGCGG-----CGCGCGGATGACGCGCGGCTG 4879
 QY 6673 TCGCGAGCGCTGCGCGCGCGCTGCGGAGCTGATGCGGAGGCGGAGCGGATCTCA 6732
 DB 4880 CTACCGGAATGATACCGGATCTGCGGCTGCTGATGATGATGATGATGATGATGAT 4939
 QY 6733 GTCAGCGCGGAGGCTGCGGAGATGCGGCTGCGGCTGCGGAGATGCGGAGATGCG 6792
 DB 4940 GTCGCGCGGAGGCTGCGGAGATGCGGCTGCGGCTGCGGAGATGCGGAGATGCG 4999
 QY 6793 GCGCGGCTTACCGGAAATTTGCGTAAAGTGCAGCATGCGGTGCGCGGAGATGGA 6852
 DB 5000 GTTTCGATCTATGCGGCGCATGCGGCTGCGGATGCGGAGGCGGCGGAGCTATCG 5059
 QY 6853 TACGCGCGAGCGGACGATGAAAGGCGGAGGCGGCTGCGGAGGCGGAGGCGG 6911
 DB 5060 CCAAGGATGACGATGCGGAGGCGGCGGAGGCGGAGGCGGAGGCGGAGGCGG 5118

RESULT 13
 ATUERTWA 5373 bp DNA linear BCT 28-AUG-2002
 LOCUS Paracoccus sp. MB1C1143 crlw, crtZ, crty, crtI, crtB genes.
 DEFINITION complete cds.
 ACCESSION D58420
 VERSION D58420.1 GI:1136630
 KEYWORDS
 SOURCE Paracoccus sp. MB1C1143
 ORGANISM Paracoccus sp. MB1C1143
 Bacteria; Proteobacteria; Alphaproteobacteria; Rhodobacteriales;
 Rhodobacteraceae; Paracoccus.
 REFERENCE 1 (bases 1 to 5373)
 AUTHORS Miki, W. and Ohnami, T., Kondo, K., Yokoyama, A., Satomi, Y., Saito, T.,
 Cantanaxanthin biosynthesis by the conversion of methylene to keto
 groups in a hydrocarbon beta-carotene by a single gene
 JOURNAL Biochem. Biophys. Res. Commun. 209 (3), 867-876 (1995)
 MEDLINE 95251715
 PUBMED 7733978
 REFERENCE 2 (bases 1 to 5373)

AUTHORS Misawa, N., Satomi, Y., Kondo, K., Yokoyama, A., Kajiwara, S., Satou, T., Ohtani, T., and Miki, W.
TITLE Structure and functional analysis of a marine bacterial carotenoid biosynthesis gene cluster and astaxanthin biosynthetic pathway
JOURNAL J. Bacteriol. 177 (22), 6575-6584 (1995)
MEDLINE 96062243
PUBMED 7592436
REFERENCE 3 (bases 1 to 5373)
AUTHORS Misawa, N.
TITLE Direct Submission
JOURNAL Submitted (23-JUN-1995) Northhiko Misawa, Marine Biotechnology Institute; 3-75-1, Heita, Kamaishi-shi, Iwate 026-0001, Japan
(E-mail:northhiko.misawa@bio.jp, Tel:81-193-26-6581, Fax:81-193-26-6584)
FEATURES
source Location/Qualifiers
1..5373
/organism="Paracoccus sp. MBIC1143"
/mol_type="Genomic DNA"
/db_xref="taxon:81397"
/note="Agrobacterium aurantiacum is distributed from Marine Biotechnology Institute as Paracoccus sp. MBIC1143. Its 16S rRNA sequence is shown in accession AB008114.-cloning vector: PAK32"
139..867
/gene="crtW"
139..867
/gene="crtW"
/citation=[2]
/codon_start=1
/transl_table=1
/product="beta-Carotene ketolase"
/protein_id="BAA09591.1"
/db_xref="GI:1136631"
/translation="MSAHLPRKADLTATSLIVSGGIIAAMLAHVAHMFDAADHPI LAIAPFLGTLTWLTVGLPIAHADMGSVPRPANAQOLVIMLVAGFSWRMTYK MAHHRAGTDDDDPDHGGVRYWAFIGYFGMRGLLPVITVYALILGDRMY VFWPLPSILASIQLFVGTWLPFRPGHDAVPPDHANRSSISDPVSLNCFHGGYH VEHHLAPTVWWRLLPSTRTKQDTA"
864..1352
/gene="crtZ"
864..1352
/gene="crtZ"
/citation=[2]
/codon_start=1
/transl_table=1
/product="beta-Carotene hydroxylase"
/protein_id="BAA09592.1"
/db_xref="GI:1136632"
/translation="MTNPLIVATVLMVLTAYSVHRMIMHGPLGWMKSHHEBHDH ALEKNDLYGLVFAVATVTLPTVGMITMAVLMVIALGMVTVGLIYFVLHDDGLVHORMP RYIPRKGYARRLYQAHRLHNAHVEGRDHCVSFGFIYAPVDKLDKMSGVLRAEAS RT"
1349..2509
/gene="crtY"
1349..2509
/gene="crtY"
/citation=[2]
/codon_start=1
/transl_table=1
/product="Lycopene cyclase"
/protein_id="BAA09593.1"
/db_xref="GI:1136633"
/translation="WTHDVLVLAGLNGLIALLARAPDLRVLLDDHAAGSDGHT WSCHPDLSPLMLRLKPLRANPDDVPRPHARRIATYGSIDGAALDAVVRG ABEIRMDSIDALDQAGATLSCGTRIEAGAVLDGGAQPSRLTVGPKFVGELETBR AGVPRPMIMATVTOODGYRFIYLFPSPRIILIDTRVSDGDDDLAAASHYV AROOGMTGAEBVRERGIPLIALADHAFMDHAGAPVPLRAGFPHPVTVGSLPAA AOVADVAGLSGPGCTALRGATIDVAIDRARRDPLRLNRMIFRCGADPRRYTLIQ RFTMPHGLIRFYAGRLSVADQLRITVTKRPITPLGTAIRCLPRPLIKENA"
2506..4011
/gene="crtI"
2506..4011
/gene="crtI"
CDS

/gene="crtI"
/citation=[2]
/transl_table=1
/product="Phytoene desaturase"
/protein_id="BAA09594.1"
/db_xref="GI:1136634"
/translation="MNAHSPAAKTAIVIGAGFGGLALAIRLQSGAGIATTLVBARDKRG GRAYVNDQDGLDAGPTVITDPPDLALKEMLALGQDMARDVTLMPSPEYRLMMPGSK VPVYVNEADOLEROIQAFNPDDLEGRVPRDVAEEYVGGVYLGTVPLKLGOMLKA APALMLKATKSYHAKVATPIIDPYLRQAFSHYTLVNGNPSTSYALNHLBERG GWVFAAGGTQVLVAGVALFERDGGMLNAKVARIDTGGPRATGTVTLADGRALYDM VASNGVMNVNDRDLGHTARQSRASLNKRWMSLFWLHFGRLAPADVDVAHTLLE GPYKELVNEIFKSPILADPFLSYLHSPCTPEMAMPFGSTHYVLAAPVPHLCRADID MAYEGRYVADRILASLEERLIPYLRLANLFTTRIFTPSDPASELNHSGASFVEPILT QSWPRPHNRDKTIRNFYLVGAGTHHGAGIPGVYSKAKTAQVMSDLASA"
4008..4913
/gene="crtB"
4008..4913
/gene="crtB"
/citation=[2]
/codon_start=1
/transl_table=1
/product="Phytoene synthase"
/protein_id="BAA09595.1"
/db_xref="GI:1136635"
/translation="MSDVLVLTSTREATTQSGSPATAKLMPGIRDITVWLVMCRHA DPVYDQALGSCPEAVNDDPQARLDGRADTLALOGDGPVTPPACAVARRHDFQAM PMDLISFAMDVBARPYRLIDVLEYSYHVAIGVYGMARWVKYRDDPYLDRACTGL AFQLTINARDVIDDARIGRCYLPGDMLDQAGARVDGVPSPBLVITVILRLDAELLY ASARVGLADLPKRCWMSIAAALRIYRAILGIRLRKGGPEAYRORISTSKAKIGLIG GMDVADHACGGRVATGTLWTRPHHA"
ORIGIN
Query Match 12.1%; Score 1053.6; DB 1; Length 5373;
Best Local Similarity 57.7%; Pred. No. 1,5e-187;
Matches 2079; Conservative 0; Mismatches 1469; Indels 54; Gaps 9;
QY 3334 CTGAGCGCGCGCGCTGATTTGTCGACGAGCGCTGTGCACGACGAATCTGCTGCG 3393
DB 1271 CTATGCGCCCCCGGTGCAAGCTGAGACGACCTGMAAGATGTCGGCGCTGCGCGGC 1330
QY 3394 GGAAGCCACCTGATGCGCGCACCTTATGATGTCATTTGTCGTCGTCGCGCTGCGCTAAC 3453
DB 1331 GGAAGCGCAGAGGCGACGTCGACCATGACGTCGTCGCGCAGGCGCGCTTGCCAAC 1390
QY 3454 GGGCTGATTCGCTGCTGCTTTACGCCGACGCTGACCCGCACTTAAAGTTTGTACTAGAG 3513
DB 1391 GGGCTGATTCGCTGCTGCTGCGCGCGCGCGCGCGCGCGCGCGCTGCTGCTGAGAC 1450
QY 3514 AGTCAGGCGAGCGCGCGCGCAATCATCTGTCGTTCCATCGCGAAGAGCTCAGCGAA 3573
DB 1451 CATGCCGACGACCGCTACAGCGCCACCTGATCTCGCCACGACCCCGACCTGTGCCCG 1510
QY 3574 GCGCAGTTTTCGCTGCGCTGAGCGCGCTGTTGCGCGCGCTGCGCCGCTTATCAGGTACGC 3633
DB 1511 GACTGCTGCGCGCGCTGAGACCCCTGCGCGCGCGCAACTGCGCCGACGAGAGTGC 1570
QY 3634 TTCCCCACCTGCGCTGCGCGCTGAGTGTGAATATTGCTTCGANTTGGCTCGGAGANTTT 3693
DB 1571 TTTCCTCCGCGCATGCGCGCGCTGCGCACCGGTTACGGGTGCTGAGCGGCGCGCGCTG 1630
QY 3694 GCCCGGCACTTACAGCAGTGTCTCGGTCGCGCGCGCTACGACCCGAGCGCGCTCAGCGAG 3753
DB 1631 GC--GGATGCGGTGTCGCGCTGCGGCGCGCGAATCCGCTGGGACAGCGACATCGCCCTG 1687
QY 3754 GTCTCACCCACCGGCGGTCAAGTGTGCGGATGCGCGGATGTTTACAGCGCAGCGCGTGATT 3813
DB 1688 CTGATGTCGAGAGGCGGAGCGCTGTCTGCGGACCCGAGATCAGAGGCGGCGCGTCTG 1747
QY 3814 GACGAGCGCGGCTGCGAGCGCGCAACCGCATCTGCACTGCGGTATCAGGATTTGTCGGT 3873
DB 1748 GACGCGGCGGCGCGCGCGCTGCGCGCATCTGACCGTGGGTTTCCAGAAATTCGTGGGT 1807

QY	3874	CAGAGTGGCAACTGGCCGCGCCGCAATGGCCTCGACAGACCCATATATTGATGATCGCAAGC	3933
Db	1808	GTCCAGATTCGAGACCGGACCGGCCCCACCGGCGTGGCCCGCCGACGATGATCATGAGCGACAC	1867
QY	3934	GTGATCAGACGACGGGTTATGCGTTTGTTTAAACCCGCGCCGTCAGTGGCCAGCGCGTTTA	3993
Db	1868	GTCAACCAAGCAGGACGGGTACCGCTTCACTATCTGCGCCCTTCTCTCGACCGCATC	1927
QY	3994	CTGATTGAGATACCCACTATCATCAACCATGCAACGCTGATGTGCGGACAGGCGCGCGT	4053
Db	1928	CTGATTCAGAGACACGCGCTATTCCGATGCGGCGCATCTGACACGACGACGCGTGGCGCG	1987
QY	4054	CACATTACGATTAATGCCACACGCGCGGCTGGAATTTGGCCGACGCTGTGCGGAGAG	4113
Db	1988	GCATCCACGACTATATGCCCGCCACACAGGAGCTGACACCGGAGCGCGAGTCCGCGC---GAA	2044
QY	4114	CAGGCTTCGCTGCGCGATACGCTCATAGCGGCGATATCATGTTCTGG---CAACAGCAG	4170
Db	2045	CGCGGCACTCTTCGCATGCGGCTGAGCCCATATGCGGCGGAGCTTTGTGGCCGATCAGCG	2104
QY	4171	CACGGGCAACCGTGCAGAGGCGGCTGCGGCGCGGACGTTTTCAGCGCACCAACCGGTTACTCG	4230
Db	2105	GCGGGGCTGTTCCTCCGTGGAGACTGCGCGGAGGTTCTTTCAATCCGATCACCGGCTATTGC	2164
QY	4231	CTGCCGCGCGGATGGCGCTGCGGCGGAGAAATTGCAAG---CAGCGTCCCGCGGACGCT	4287
Db	2165	CTGCCCTTATGGGCGACAGGTGGCGGAGCGTGTGGGGGTCTGTCCGGGCGCGCCGCGCAC	2224
QY	4288	CACACGCTGAGCCACTGCATGAAATCCTTGGCCGTCAGCACTGGCGGAGCAGCGCTTT	4347
Db	2225	GACCGGCTGCGCGGCGGCATCCGAGATTACCGCATCGACCGGCGCGCGCGCATCTTT	2284
QY	4348	TTCCGTCGTAAATATCGACTGTTCTTCCGCGAGCGCCGTAAACAGCGTGGCGGTA	4407
Db	2285	CTGCGCTTTTGAACCGGATCTGTTCGCGGCTGCGCGCCGACCGGCGCTATATACCTTG	2344
QY	4408	ATGCAGCGTTTATACCGGCTTGAACGCGCGGATTTGATTAAGCCGCTTTTACGCGCGGCAACTG	4467
Db	2345	CTGCAGCGGTTCTTACCGGATGCCGATGGAATCGAATCGAACGCTTCTATGCGCGCGGCTG	2404
QY	4468	CGCTCAGCGATTAAGCAACGCAATCTGTGCGGCAAAACGCGGTCCTCTTGGCGAAAGC	4527
Db	2405	AGCGTGGGGAATCAGCTGCGCATGTAACCGGCAAGCCTCCATTTCCCTTGGACAGGCC	2464
QY	4528	CTGGCGCGATGATGATGACTCTCTCCGTACAGAGGAAGAAATTA-----	4572
Db	2465	ATCGCTGCGCTGCCCGACGTCGTCCTCGTGAAGAAACCGCATACGCCATTGCGCCG	2524
QY	4573	---TGAAACGCAACCTATGATGATTTGGGCGCAGAGCTTCGGGCGCTGGCGATTCGTC	4629
Db	2525	CGGCGMAAGCCGCATCTGATTCGGCGAGGCTTTTGGCGGAGCTGACCTTGGCATTCGCGC	2584
QY	4630	TGCAAGCGGCGCGGCGTGCCTGCTACGCTGCTGGAACAGCGCATTAAGCTTGGCGGCGCG	4689
Db	2585	TGCAGTCCGCGGGGATGCGCACACGCTGTCGAGGCCCGGAGACAGCGCGGCGGGGCG	2644
QY	4690	CTTATGTGATACAGGATCAGGGTTTACCTTTGATGCGGTTCGACGCTATACCGATC	4749
Db	2645	CCATATGTCGACGATCAGGGCCATCTTTCGACGCGGCGCCACCGTCACTCACCGAC	2704
QY	4750	CCAGGCGCATGAGGCGCGCTTTTACGCGTGGCAGGACAGCACTCAGTATATTATGCGAC	4809
Db	2705	CCGATGCGCTAAAGAGCTGTGGGCTCTGCACCGGCGAGAACATGGCGCGCATGTGACGC	2764
QY	4810	TGATGCGCGTACGCACTTTATCGCTGTGCTGGGAAGACGGCAGCGCAGCTGACTACG	4869
Db	2765	TGATGCGCGTCTCGCCCTTCTTATCGCGTGAATGTGCGCGGCGGAGAGGTCTTCGATTACG	2824
QY	4870	ACAACATCAGGCGCACTGGACGACGATGTCCACTTTTAATCCAGGATGTCCGC	4929
Db	2825	TGAACGAGGCGGATCAGCTGGAGCGCCGATGCGCCAGTTTCAACCCGAGCACTGGAGG	2884

QY	4930	GTAAACCCGCAATTTCTGGCTATTCAAGATGTGTTTCGTGAAGGCTATCTGAAACTGG	4969
Db	2885	GATACCCGCGCTTCCGTGATTACCTAGAGGAGGTATCACAGAGAGGCTACGTCAAGCTGG	2944
QY	4990	GCACCGTACCTTTCTGCAATTTCCGGACAATAGCTGCGGTGCGGGCCACAGCTGGATCGGC	5049
Db	2945	GCACCGTGCCTTCTCTCAAGCTGGGCGAGATGCTCAAGGCCGCGCCGCTGTATGAAGT	3004
QY	5050	TGCAGGCTGCGCAGATGCTTCAACAGCATGTGTGGCAATTTATTCATGACGATCATCTGC	5109
Db	3005	TGGAGGCTTAACAATCGGTTCATCCCAAGGTCCGACCTTCATCAAGAACCCCTATCTGC	3064
QY	5110	GCCAGGCTTTTCTCTTCACTCGTTGCTGGTCCGCGGTAATCTTTTGGCAAGCTTTTCA	5169
Db	3065	GGCAGGCTTTTCGTATCACACGCTGTGTGGCGGGAAATCCCTTCTCGACCAAGCTCGA	3124
QY	5170	TCTATACCTTAATTCAGCAGCTGGAAGGCGCAATGGGGCGTGTGGTTTCCGCGCGCGCTA	5229
Db	3125	TCTATGCGCTTAACACAGCGCTGGAAGCGCGCGGAGTCTGGTTGCCAAAGGCGAGCA	3184
QY	5230	CCGGTGCCTGTTGATATGACATGCGCGCGGCTGTTTCGCAATTTGGGCGGTAACTGCTGC	5289
Db	3185	CCAAACAGCTGTTGCGGCGCATGTGTCCGCTGTTTGCAAACGCGCTTGGGGGAGATGTCTGC	3244
QY	5290	TCAAACCCGAAGTACGCCAGCTGGAAGACCGAGGTTAACCGCATAGCGGTGTCCAGCTGA	5349
Db	3245	TAAACGCCAAGTACGCGCGGATTCGACACGAGCGGGCCGCGCGACCGGCGTACCTGTG	3304
QY	5350	AGGATGGGCGCGGTTTGTGCGCGCCGCGCTGTGGCTCAATGCTGAAGTGTGATACCT	5409
Db	3305	CCGACGGGCGGCGCTTGAACCGCCGCAATGTGTGCGACCAACGGGAGAGTGAATGCACACT	3364
QY	5410	ACGATCGCTGTTAAGCCAGCATCTCGCGGCGCGCTAAACGCGCGGCAACGCTGAAGGCA	5469
Db	3365	ATCCGGAATCTCTGGGCGCATACCGCCCGCGGGCAGAGCCGGGCGAAATCTGTCGAACGCA	3424
QY	5470	AGCGATGAGCAATCGCTGTTTGTACTTAATTTGGTCTTAAATCATGCCACCCGAGC	5529
Db	3425	AGCGGTGTCATGTGCGCTCTTGTGTGATTTCCGCTGCGGAGGCGGCCAAGAGACG	3484
QY	5530	TGGCGCACCAACGCGTGTGCTTGTGTCCGCGCTATCGTGAATTGATGAGATCTTCA	5589
Db	3485	TGGCGATCAACACATCTGTTCGCGCCCGCGCTCAAGAGATGTGTCAACGAGATCTTCA	3544
QY	5590	ATAGAGCACACTGCGCGGGAATTTCTCGCTATCTGCAATGCGCCCTGCGCCAGCATC	5649
Db	3545	AGGCGCGAATCTGCGCGAGATTTCTGCTCTATCTGCAATTCGCCCTTGACACGACAC	3604
QY	5650	CGTGGCTGGCACCGGCGGCGCTGCGGAGTTTAACTGTGCGCGCGGATGCGCGATCTCG	5709
Db	3605	CGAGATGGCGCTCCGGGCAATGTCACAGCATTAAGTCTGTGGCCCGGTGCGCATCTGG	3664
QY	5710	GTACCGCGCAATTGACTGTCGCAACAGAAAGGCGCGCTTGCGCGATTCGATTTTGCTT	5769
Db	3665	GCCGCGCGAATGATTGGCGGCTGCAAGGGGCGCGCTATGCGGACCGCATCTTGCGCA	3724
QY	5770	ATCTGAGAGACACTATATGCGCGGCTGCGGACAGCAGTATAGTACACACCGTATGTTTA	5829
Db	3725	GCTTGAGAGACGCGCTGATCCCGAATCGCGCGCAACTGACCAACAGACGACATTTCA	3784
QY	5830	CGCGCTTGAATTTTCGACACACGCTGACAGCGCATCAAGGCTCAGCGTTTTCGCTCGAAC	5889
Db	3785	CCCGTCCGATTTTGGCAGAGGAATGAACGCGCATCATGGAAGGCGCTTCCGGTGCAGC	3844
QY	5890	CCATTTTGAAGCAAGCGCTGTGTTCCGGCCGCAATAACCGGATGCGCAATTACTTAAC	5949
Db	3845	CGATCTGACCAATCCGCGTGTGTTCCGGCGCGCAACCGCGACAMAGCATCCGCAACT	3904
QY	5950	TTTATCTGAGTGGGGCTGCGACGATCCCGGTGCGGGTGTGCGAGGGGTATATCGGCTCCG	6009
Db	3905	TCTATCTGAGTGGGGCGGAGCAACCATCCGGGCGGGGGAATTCGCGGCGTGTGTGGGCTCGG	3964
QY	6010	CGAAAGCAGACCGCCAGCTGATGTGTGAGGATCTGACCGGATGAACCAACCGCCGCTGAT	6069

[illegible]

SOURCE	Flavobacterium sp. ATCC 21588
ORGANISM	Flavobacterium sp. ATCC 21588
	Bacteria; Bacteroidetes; Flavobacteria; Flavobacteriales;
	Flavobacteriaceae; Flavobacterium.
REFERENCE	1 (bases 1 to 5188)
AUTHORS	Passamontes,L., Hug,D., Teasler,M., Hohmann,H.P., Schierle,J. and van Loon,A.P.
TITLE	Isolation and characterization of the carotenoid biosynthesis genes of Flavobacterium sp. strain R1534
JOURNAL	Gene 185 (1), 35-41 (1997)
MEDLINE	97186694
PUBMED	9034310
REFERENCE	2 (bases 1 to 5188)
AUTHORS	Passamontes,L.
TITLE	Direct Submission
JOURNAL	Submitted (03-JUL-1996) VFCB, F. Hoffmann-La Roche Ltd, Basel, 4070, Switzerland
FEATURES	Location/Qualifiers
SOURCE	1..5188

RESULT	14		
FAU62808/c			
LOCUS	5188 bp	DNA	linear
DEFINITION	Flavobacterium ATCC31588 geranylgeranyl synthase (crtE), phytoene synthase (crtB), phytoene desaturase (crtI), lycopene cyclase (crtJ), D-carotene hydroxylase (crtZ) genes, complete cds.		
ACCESSION	FAU62808		
VERSION	U62808.1	GI:1842241	
KEYWORDS			

gene

terium sp. ATCC 21588
terium sp. ATCC 21588
; Bacteroidetes; Flavobacteria; Flavobacteriales;
teriacae; Flavobacterium.
s 1 to 5188)
es, L., Hug, D., Tesser, M., Hohmann, H. P., Schierle, J. and
n and characterization of the carotenoid biosynthesis genes
bacterium sp. strain RJ534
(1), 35-41 (1997)

es 1 to 5188)
es, L.
ubmission
d (03-JUL-1996) VPCB, F. Hoffmann-La Roche Ltd., Basel,
tzerland
Location/Qualifiers
1. 5188
/organism="Flavobacterium sp. ATCC 21588"
/mol_type="genomic DNA"
/strain="RJ534"
/db_xref="taxon:50286"
95. .982
/gene="crte"
95. .982
/gene="crte"
/codon_start=1
/transl_table=11
/product="geranylgeranyl synthase"
/protein_id="AAC44848.1"
/db_xref="GI:1842242"
/translation="WTPKQFPRLDVLVETRLAQISQGEVSNAPLGAANSDALSPK
RFRVALVIMVAVSSGVCDAWDAACVAMVHAASLI FDDMPCMDARTTRGDPATV
AHGEBAVYAGIALITTEMYLIGGEARPPDPRARIVASMRAMGVGICAGDLDLH
APKDAAGIEBDDLTGVLFVFGLEMLSTIKGLDKAETSOLMAFGRLFGVSYDIL
LDVIGDKRSTGDDTRADVNAAPPKKGMVAVGMGVDAVGHYRASRAQLDELMLRRLPRG
GQIALDLARVLPHDTRRBA"
complement (979. .1890)
/gene="crtb"
complement (979. .1890)
/gene="crtb"
complement (979. .1890)
/codon_start=1
/transl_table=11
/product="phytoene synthase"
/protein_id="AAC44849.1"
/db_xref="GI:1842243"
/translation="MTDLTATSEAAIAAGSGSPAOKAKMPPGIRBDTVMLVAMCRHA
DDVIGVGMGSAPEAGCPBOAATGLRADTLAAIHEDGMSPPRALROVARHPDD
LMPHDLIGFAVDVADREYRSLDDVLEYSYHAGVGVGMARVMGYQDDVADRACDL
GLAFLNTLIGFAVDVADREYRSLDDVLEYSYHAGVGVGMARVMGYQDDVADRACDL
VYASRQGI.PHLPKCAWSIAALRIYRAIGTRIRGGPEAYRORISTSKAATIGLAA
RGDLAAASRLGGEISRDGLWTRBRA"
complement (1887. .3371)
/gene="crtI"
complement (1887. .3371)
/gene="crtI"
/codon_start=1
/transl_table=11
/product="phytoene desaturase"
/protein_id="AAC44850.1"
/db_xref="GI:1842244"
/translation="MSSAIVIGAGFGGLALAIRLOSAGIATTTVEARDKRGGRAYVNN
DOGFVADGAPVTPDPSIRBELMAISGQPMEDVTLLPSPFRLTMWAGRSPEYVND
DDELIRVASFPAVDVGDREHDAVEYRGGYLLGTTPFLKLGOMLAAALMALML
GAYBSVMVAFFDOPHLROASPHHTLLVGNPSTSIYALLHLEBRGVMFARKG
CTNQLVAGMVALFELGKLTLLNARVYTRIDTEGDRAITGTLIDGRDLRADTVASNGDV
MSYVDLLGHTRRGRTKALINLRQWSMLPVHLFGLSKRPENLHSHYIFGRYKGL
VNIIEFELERRLLPDPMYLHSPCVTPDPSLPEGMSTHYVLAVPHLGRADVMAEAPG
YAEIRFEELERRLLPDPMYLHSPCVTPDPSLPEGMSTHYVLAVPHLGRADVMAEAPG
HNRDALINFYVAGGTHPGAGIPGVGSAAKNAQVMSDLAVA"
complement (3368. .4516)
/gene="crtY"

CDS

complement(3368. .4516)
/gene="crtY"
/codon_start=1
/transl_table=11
/product="lycopenene cyclase"
/protein_id="AAC44851.1"
/db_xref="GI:1842245"
/translation="MSHDLLIAGAGLSGALLALAVDRRPPARIIVMLDANSGRSDQHT
WSCHDTHVLSPEWRLATSPTRRGEWTDQVAPPDHSRLITVYGSIEGVALIGLQVAD
LWMTHTVLLDDTGATLTDGSRLEAACAIDARAVETPHLTVGQKVEGALIEIDDA
GGERPMIMQATVPQMDGVRFTLLPFSPRTLLIEDTYSDDGDLDDGALLQASIDPA
RGGWGMQERREBRGTLPIALAHDAIGFMRDHAQCAVPVGGAGLPHYVTSI.PYAAO
VADALAAARLTTASARRAVRGVADIRARDFPLRLNMLFRGCPPPRRRLRLORFTR
LQOPLERRYAGRLTLADRLRITVTRPPIPLSQAVCLPERPILQERA"
complement(4513. .5022)
/gene="crtZ"
/gene="crtZ"
complement(4513. .5022)
/gene="crtZ"
/codon_start=1
/transl_table=11
/product="b-carotene hydroxylase"
/protein_id="AAC44852.1"
/db_xref="GI:1842246"
/translation="MSTAAAILTVLTITAAAMELTRSYVHRIMHGCPICMGKHSNHSDE
DHDDHLEKNDLIGVIFAVISTVLPFATIGMGSDLMWMLAVGTGGLIYFPLHDLVNG
RMPEFVPPRGRLRRVYQAHNRHNAHGRENCVSFGFTWAPSDLSLAEELKRSGLALK
DREGADRYT"

gene

CDS

ORIGIN

Query Match	11.9%	Score 1032.2;	DB 1;	Length 5189;
Best Local Similarity	57.1%;	Pred. No. 1.6e-183;		
Matches 2031;	Conservative 0;	Mismatches 1493;	Indels 33;	Gaps 7;

Qy	3419	TTATATGTCATCTCTGGTTCGTGCTCCGGGCTTGGCTTAACGGGCTGATATGGCTGCGTTTACGCC	3478
Db	4509	ATGATCTGTCTATCGCGGGCGCGGGGCTGTCCGGTGCCTGATGCGCGTTGCCGTTGGCG	4450
Qy	3479	AGCTGCAGACCCGGCACTTAAAGGTTTTTGTCTACTGAGAGTCAAGGCGAGCCGGCCGGCAATC	3538
Db	4449	ACCCGAGACCCGGATATCGCGCATCGTGAATGCTGACGCGCGGTCCGGCCCCCTGGAACCAAGC	4390
Qy	3539	ATACCTGGTCGTTCCATCGCGAAGACGTCAAGCAAGCCGACAGTTTCGTGCTCGACCGCGC	3598
Db	4389	ACACTGGTCTCTGCGACAGACAGCATCTTTGGCCCGAATGGCTGCGGCTGTGCGCCA	4330
Qy	3599	TGCTTTCCGGCGCGCTGCGCCCGGTTATACGTAACGCTTCCCAACCTGCGTCCGACGTGG	3658
Db	4329	TTCTGTCCGGCGAATGACCGGATCAGAGAGTGGCGGTTTCCCGACCATTTCCGCGCGCGCTGA	4270
Qy	3659	ATGGGAAATATTGCTCGATTCGCTCCGAGAGATTTTGGCCGGCACCTTACAGAGAGTGGCTCG	3718
Db	4269	CGACAGGCTATAGGCTTCGATCGAGGGGGGGCGGCGCTGATGGGCTGCTGAG-----G	4219
Qy	3719	GTGCGCGCGCTTACGCAACCGCACGCGCGGTGACGAGTCTCAACCAACCGGGGTCAAGACTGG	3778
Db	4218	GTGTGCAATCTGCGGTGGAAATACGATGTGCGGACCGCTGGAGATACCGCGCCGACGCTGA	4159
Qy	3779	CGGATGCGCGGATGTTACAGCGCGACGCGGTGATTAACGACGCGCGGCTGACCGCCAGAC	3838
Db	4158	CGGACCGGCTCCGGGATCGAGGCTGCTGCTGATGACGCGCCGTGGTGCCTGCGAGACC	4099
Qy	3839	CGCATCTGCACTGGGCTATACGAGCATTTGTCGTCAGGAATGGCACTGGCCCGCGCGC	3898
Db	4098	CGCACTCGAACCTGTGGGTTTCCAGAAATTCGTGGGGGTGCAATCGAAGCCAGACCCCCC	4039
Qy	3899	ATGGCTGCGACAGCAATATTGAATGACGCGCAGCGTGCATCACAGCAGGGATTATCGCT	3958
Db	4038	ATGGGCTGAGCGGCCGATGATCATTGACGCGGACCGCTTCCGCAATATGAGCGGGATACGCGT	3979
Qy	3959	TTGTTTACACCTTGCGCGCTCAGTGCAGCCGTTTACTGATTGAAGATACCACTATCATCA	4018
Db	3978	TCATCTATCTGCTGCTTCACTGCTCCACCCGCACTCTGATGAGAGATTCGGCTATACAGCG	3919

QY	4019	TCATGCCAGCTGAGATGCGCAAGAGCGCGCCCTCACTTACGGATTATGCCACAGC	4078
Db	3918	ACGGCGGGGATCTGAGACGATGCGCGCTGGCGCAGGCGTGCCTGGACTATGCGCCAGGC	3859
QY	4079	GGCGCTGGAAATTTCCGCGCAGCTGCTGCGCAGAGACACGGCTCGCTCCGATCAGCTCA	4138
Db	3858	GGGCTGTGACCGGGCAGAGATATGCGCGCGAAAGGGGATCTGCGCATCGCGCTGGGCC	3799
QY	4139	GGCGCGATATCGATCAGTTCGTGGCAACGACGACGCGGCAACCGTGCAGCGGCTGCGCG	4198
Db	3798	ATGACGCCATATGCTTCTTGCGCGACACGCGCAGGGGGCGGTGCGGTTTGCGCTGGGGGG	3739
QY	4199	CCGCACTGTTTACGCCACACCGGTTACTGCTGCTCCCGCGCGTGGCGCTGGCGGAGA	4258
Db	3738	CAGGCTGTTCACCGCGTCAACCGGATATTCGCTCCCTATGCGCGCACGGTTCGGCATG	3679
QY	4259	AGATTGCGACAGCTGCGCGCGACGCGTCAACAGCTGACGACATGCAATCGAATCTTTTG	4318
Db	3678	CCATTCGCGCGCGACCTGACACGCGCTCGCCCGTGGCGCG--TGCGCGGCTGGG	3622
QY	4319	CCCGTCAGCACTGCGCGGAGCAGCGCTTTTCCGCTGTTAAATCGCATGCTGTTCTTG	4378
Db	3621	CCATCGATCGCGGGATTCGCGACCGGCTTCTCGCGGCTGTGAACCGATGCTGTTCCGG	3562
QY	4379	CCGACCGGCTGAACAGCGCTGCGCGGTATGACGCGTTTATCCGCTGAACCCGGAT	4438
Db	3561	GCTCCCGCGCGACCGTGTGCTATGCTGCTGCGACGCGTTCCTACCGGCTCGCAGCGCG	3502
QY	4439	TGATTACCGGTTTATGCGCGGGGCACTGCGGCCCTACGCGATTAAGCAGCATTCGTGG	4498
Db	3501	TGATTCGAGCGTTTATGCGCGGGGCGCTGACATTTGCGCAGCGGCTTGACGCTACCG	3442
QY	4499	GCAACGCGCGGTCCTCTCGGCGCAACGCGCTGCGCATTTGATGATGACCTTCGCTTAC	4558
Db	3441	GACCGCGCGCATTCGCGTGCAGGCGCGTGCCTGCGCGAAGCGCCCTGCTGCG	3382
QY	4559	CAGGAAAGAAATTAAGAAAGCACTTATGATTTGCGCAGGCTTCGCTGCGCTGCGCT	4618
Db	3381	--AGGAGAGAGCATGATTCGCGCATTCGCTGATGCGCGCAGGTTTCGGCGGCTTGGCGCT	3325
QY	4619	GGCGATTCGTTGGAACCGGCGCGCGCGCGCTGACGCTGCTGAAACAGCGCGATAGCC	4678
Db	3324	TGCCATCCGCTGGAAATCGCGCGGATTCGGACACCATGTCAGAGCGCGCGCAAGCC	3265
QY	4679	TGGCGGGCGGCTCATGATGATCAGAAACAGGGTTTAACTTTATGCGCGTCCGACGCT	4738
Db	3264	CGGCGCGCGGCTTATGTCGGAACATCAAGGCGCACAGTCTTCATGACAGGCCCGACGCT	3205
QY	4739	GATTACCGATCCACGCGCTATCGAGGCGCTGTTTACGCTGGCAGGCAAGCACTCAGTGA	4798
Db	3204	CGTACCGGACCCGACAGCGCTGCGAGAGCTGTGGGCCCTCAGCGGCGCAACGATGAGCG	3145
QY	4799	TTATGTGACCTGATGCCGCTGACGCCATTTTATCGCTGTGCTGGGAAGAAGCGCAGCA	4858
Db	3144	TGAGGTGACGCTGCTGCGGCTTCGCGCTTCACCGCTGACATGCGCGGACGCGCCAG	3085
QY	4859	GCTGACCTAGCACAACATCAGAGCGCAGCTGGAGACAGATTGSCACTTTTATCCCA	4918
Db	3084	CTTGAAATACGTGAACAGACGACGAGCTGATTCGCAAGGTCGCTTCCTTCAATCCGC	3025
QY	4919	GGATGTGCGCGTTACCGCAGTTTCTGAGCTTATTCACAGATGTGTTTGGTGAAGGCTA	4978
Db	3024	CGATGTGATGAGCTATGCGCGCTTCCAGATTAAGCGGAGAGGTCTATGCGCAGGGGTA	2965
QY	4979	TCGAAAACCTGGGACCGTACCTTTTCTGCAATTCGCGACATGCTGCTGCGGCGCAC	5038
Db	2964	TCTGAACCTGGGGACACGCGCTTCTGAAAGCTGGGCGCAATGCTGAACCCGCGCGGC	2905
QY	5039	GCTGGGTGCGTGAAGGCGCTGGCGCAATGTCATCAGATGAGTGGCGAAATTTATTCGA	5098
Db	2904	GCTATGCGCTGAGGACATACGCTCGGTTCACAGATGATGTGGCGGCTTCAACGGA	2845
QY	5099	CGATCATCTGCGCAGGCTTTTCTTCTTCACTGCTGCTGATCGGCGGTATCCTTTTGC	5158

Db 2844 CCCGATCTCGGAGGCGCTTCTGTTCCACACGCTGCTGCTGCGGGAACCCGTTTC 2785
Qy 5159 AACGCTTCGATCTAATACCTTAATTCAAGCACTGAGCGGCAATGAGGCGGTGTTTC 5218
Db 2784 GACGAGCTCGATCTAATGCGCTGATTCATGCGCTGGAAAGGCGGCGCGTCTGTTCCG 2725
Qy 5219 GCGGCGGCTAAGCGGCTGCTGTTGATGCGATGCGCGGCTGTTTGCGATTTGGCGG 5278
Db 2724 CAAGGCGGCGACCAACAGCTGCTGCGGCGATGTCGCCCTGTTCCAGCGCTTGGCGG 2665
Qy 5279 TGAATGCTGCTCAACGCCAAGTCAAGCAGCTGAGAGCGAGGTTAACCGCATCAGCGG 5338
Db 2664 CAGCTGCTGATGATGCGCGCTGACGCGGATGACACGAGGCGCATGCGCGACAGG 2605
Qy 5339 TGTCCAGCTGAAGATGAGCGCGCTTGTTCGCGCGCGCGCTTGCCTAATGCTGACGT 5398
Db 2604 CGTACGCTGCTGATGACGCGGCGGAGTTGCGCGGATACGCTGCGCAAGAACGCGACGT 2545
Qy 5399 GGTGCAATACCTAAGATGCGCTGTTAAGCAGCATCTGCGGCGCTTAAAGCGCGCAC 5458
Db 2544 GATGACAGCTATGCGCGACCTGCTGCGCATACCGCGCGCGGCGACCAAGCGCGCAT 2485
Qy 5459 GCTAAGCGCAAGGAGATGACCACTGCGCTGTTGTACTAATTTGGCTTATCATGC 5518
Db 2484 CTTAAGCGGAGCGCTGCTGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2425
Qy 5519 CCACCCGAGCTGCGGACCAACACGCTGCTTGTTCGCGCGCTATCGTGAATGATCGA 5578
Db 2424 CCGGAGAACCTGGCCACCAAGAGTCATCTTCCGCGCGGCTACAAAGGCGCTGATGA 2365
Qy 5579 TGAATCTTCAATAGCAAGCTGCGGGAAGATTTCTGCTGATCTGCAATGCGCGCTG 5638
Db 2364 CGAATCTTCAACGCGGCGACGCTGCGGAGATTTCTGATGATCTGCAATGCGCGCTG 2305
Qy 5639 CTCGAGCATCCGTCGCTGCGACCGCGGCTGCGGCGAGTTTAAAGCTGCGCGCGT 5698
Db 2304 CGTGACGATCCAGCTGCGCGCGCGCGGAGATGCCAGCATTAACGCTTGGCGCGT 2245
Qy 5699 GCCGATCTCGGTAACCGCGCAATTGACTGCGCAACAGGAAGGCGCGCTTGGCGGATCG 5758
Db 2244 TCCGATCTGGCGCGCGCGATGTCGATTTGGGAAGCGAGGCGCGGCTATGCGGACG 2185
Qy 5759 CATTTTCTTATCTGAGGAGCACTAATGCGCGGCTTGGGACAGCAAGTAAAGTACCA 5818
Db 2184 CATCTTCAGGAACCTGAGCGCGCGCGCATCCGACCTGCGCAAGCACTGACCGCTAG 2125
Qy 5819 CCGATGTTTAAAGCGGTTGATTTTGGGACAGCGCTGCGACCGCATCAGGCTCAGCGTT 5878
Db 2124 CCGCATCTTACGCGCGCGGATTTTCAAGCAAGCACTGCGCGCATCAAGCGCGCGCTT 2065
Qy 5879 TTCGTCGAGCAACCATTTTGAAGCAAGCGCTGCTGCTGCGCGCAATTAACCGCATGCGGA 5938
Db 2064 CTCGATCGAGCGGATCTGAGCGCAATCGCTGCTGCTGCGCGCAATTAACCGCGCGC 2005
Qy 5939 CATTAATACTTTATCTGCTGCGGCGCTGCGACGATCCGCTGCGGCTGCGGCGCT 5998
Db 2004 GATCCCGAATTTTCAATGCTGCGGCGGCGGCAAGCATCGGCTGCGGCGCATCCGCGGT 1945
Qy 5999 GATCGGCTCGGAAAGGAGCGCGCGCATGATGCTGAGAGATCTGACCGGATGAACCA 6058
Db 1944 CGTTGGGAGCGGCAAGCGCGCGCATGCTGCTGCGGCGCGGCGGCGGCGGCGGCGG 1888
Qy 6059 CCGCGCGTGAATGAGAGGATCAAGCAACATGCGCGAGGCGCTCAAAAGTTTGGCGGCG 6118
Db 1887 ACCGATCTGAGCGGCACTTCCGAAGCGCGCATGCGCGAGGCTTCCGAAGCTTCCGCGAG 1828
Qy 6119 GCTACCGGCTATTTGATCTTCAAGCGGCGCGAGTACGCTGATGCTGACCGCTGCTGCT 6178
Db 1827 GCGGCAAGAGTATCGCGCGCGCATCGCGAGGATGAGGCTTATGCTGCTGCTGCTGCT 1768
Qy 6179 CGTCACTGTGAAGATGATAGATGTCAGAGCTGCGGCGAAAGCGCGGACGCGACGCGG 6238

Db 1767 AGGATGCGGATGACGCTGATTCAGCGGCGAGTATGCTTCCGCCGAG-----GCG 1714
Qy 6239 GTGCGGATGACAGCGCGGATGCGCACCTGCAATGGAACCGCGCGCTTACAG 6298
Db 1713 GCGGCGGACCAAGCGCGGCTGCGGCGGCGCTGCGCGGACAGCTGCGCGCTGCGAC 1654
Qy 6299 GGTGCCCAATGATGAACAGCGCTTTCGCTTCAAGAAATGCGCTGACGATCAG 6358
Db 1653 GAGAGCGGCGGATGCTGCGCGCTTTCGCGGCGCTGCGCGGCTGCGCGGCTGCGG 1594
Qy 6359 CTTCGCGAGCGCTGCTTTCATCTGAAAGGTTTTCGATGATGCGCGCTGAGAA 6418
Db 1593 TTCCCGGACTTTCGCGGATGAGCTGATGAGAGTTTCGCGATGATGCGCGGATTCG 1534
Qy 6419 CTTATGCGCTTTCGCGGACAGCTGCTGATCTGCTTACAGTTCGCGCGCGGCGGCGG 6478
Db 1533 GAATACCGGAGCTGATGACGCTGGAATATTCCTACACGCTGCGGCGGCTGCGG 1474
Qy 6479 TTAATGAGCGCGCGGATGAGGCTGATGAGCGGCTGATGAGCGGCTGATGAT 6538
Db 1473 GTATGATGCGCGGCTGATGAGGCTGAGAGAGATGCGGCTGATGAGCGGCTGAT 1414
Qy 6539 TTGAGTCTGCGCTTTCAGCTTACCAATATGCAAGGATATCTGTAAGACGCGGAGAT 6598
Db 1413 CTGGGCTTTCGCTTTCAGCTGAGCAATGCTGCGAGCTGATGAGCGGCTGATG 1354
Qy 6599 GCGCGCTTTCATCTGCGCAAACTGCTGATGATGAGCGGCTGATGAGCGGCTGAT 6658
Db 1353 GCGCGCTTTCATCTGCTGCGCACTGCTGCGGCTGCGGCGGCGG-----GCGAGGCT 1303
Qy 6659 GCGGATCCGCAATGCGCGGCGGCTGCGGCGGCTGCGGCGGCTGCTGCTGCTGCGG 6718
Db 1302 GAGGCTCGGCTGCTTTCAGCGGCTGAGCGGCTTATTCGCTATATATCCGCTGCTGAGCGGCG 1243
Qy 6719 GAGCGCTATATGATGACGCGGAGCGGCTGCGGAGATGCGGCTGCTGCTGCTG 6778
Db 1242 GAGCGCTATATGCTTTCGCGGCGGCGGCGGCTTTCGCTATGCTGCGGCGGCTG 1183
Qy 6779 GCGATGCGCACCGCGCGGCGGCTTTCGCGGAAATTTGCGCTAAAGTGCATGCGGCT 6838
Db 1182 TCAATGCGCGCGGCTGCTTATCTATGCGCAATGCGGAGCGGCTGCGGAGGCTG 1123
Qy 6839 GCCGCGGATGAGGATGCGCGGCGGCGGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 6898
Db 1122 CCGGAGGCTATGCGGCGGCGGATGAGCAAGTCAAGAGCTGCAAGATCGGCTTTCGCG 1063
Qy 6899 AAAGTTCGCGGCTGCGCTTACTTTCGCGCTTTCATCCGAGGCGGCTTTCGCGG 6958
Db 1062 CCGGAGGCTTTCGAGCGGCGGCGGATGCGCTGCGGCGGCGGCGGCAATCAAGCGGCG 1003
Qy 6959 CTGTCGAGGCTTTCGCGG 6975
Db 1002 CTGTCGAGCGGCGGCGG 986

RESULT 15
AR169831/c 8625 bp DNA linear PAT 17-DEC-2001
LOCUS AR169831 Sequence 1 from patent US 6291204.
DEFINITION AR169831
ACCESSION AR169831.1 GI:17907751
VERSION AR169831.1
KEYWORDS
SOURCE
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 8625)
AUTHORS Pasamontes,L. and Tsygankov,Y.
TITLE Fermentative carotenoid production
JOURNAL Patent: US 6291204-A 18-SEP-2001;
FEATURES
1. 8625
/organism="unknown"
/mol_type="unassigned DNA"

QY 5519 CCACCCGAGCTGGCGACCAACAGGTGTGTTTGTTCGGCGTATGCTGATGATCGA 5578
 Db 4850 CCCCGAAGACCTGGACCCACACAGCGTCATCTTGCGCCGCGCTAACAGGGCGTGA 4791
 QY 5579 TGAGATCTTCAATGAGCAGCAGCTGGCGAGATTTTCGCTGATCTGATGCGCCCTG 5638
 Db 4790 CGAATCTTCAAGGGGCGACCGCTGCGGACGATTTTCGATGATCTGCACTTGGCCCTG 4731
 QY 5639 CTCAGCGATCCGTCCTGGGACCGCGGGGCTGGCGAGTTTACGTCGCGCCGCT 5698
 Db 4730 CGTAGCGATCCAGCTGGCGCCCGAGGGGATGCCACGATTAACGTCCTTGCGCCGCT 4671
 QY 5699 GCGCATCTCGGTACCGCGCGCAATTGACTGGCAACAGAGAGCGCGCTGGCGCGATCG 5758
 Db 4670 TCCGATCTGGGCGCGCGGATGATGATGGGAAAGCCAGGCGCCGCGCTATGCGGAGCG 4611
 QY 5759 CATTTTGTCTATCTGAGAGAGCACTATATCCGGGTCGCGACAGCAGTTAGTGAACA 5818
 Db 4610 CATCTTGAAGAACTGAGAGCGCGCGCATCTCCGACCTGCGCAAGCACTGACCGCTGAG 4551
 QY 5819 CCGTATGTTTACGCGCTTGTATTTTGGGACAGCTGACCGCATCAAGGCTGACGCTT 5878
 Db 4550 CCGCATCTTACGCGCGCGCGGATTTACGACCGAACTGCGGCCCATCAAGCGCAGCGCTT 4491
 QY 5879 TTGCGTGAAGCCATTTTGAAGCAAGCGCTGCTGCTGGCGCGATTAACCGCGATGCGA 5938
 Db 4490 CTCGATCGAGCGGATCTCGAGCAATCGCTGTGCTCGCGCGATTAACCGCGAGCGGCG 4431
 QY 5939 CATTTACTAACCCTTATCTGCTGGGGGGCTGCGACCGCATCCCGGTCGCGAGCGT 5998
 Db 4430 GATCCCGAATCTTCAATCGTGGGGGGCGGCGACCGCATCCGGGTGCGGCGGT 4371
 QY 5999 GATCGGCTCCGCGAAGCGACCGCCAGCTGATGATGAGGATCTGAACCGATGAACCA 6058
 Db 4370 CGTTGGAGCGCCAGGCGACGCGGAGGTATGCTGCGGACCTGGCGG---TCGCATG 4314
 QY 6059 CCGCGCGTGAATGAGCAGGTACGCGAAACCATGCGCGAGGGCTCCAAAGTTTGGCCAGC 6118
 Db 4313 ACCGATCGAGGGGACTTCCGAAAGCGGCGCATCGCGAGGGTTGCAAAAGCTTGGCGGAG 4254
 QY 6119 GCTACCGGCTATTTGATCTTTCAACGCGCGCGCAGTACGCTGATGCTGTACGCTGTGT 6178
 Db 4253 GCGGCGAAGCTGATGCGCGCGCGCATCCGCGAGATACGTCATGCTATGCTGTGTGC 4194
 QY 6179 CGTACTGTGAAGTGTATGATGATGATGATGATGATGATGATGATGATGATGATGATG 6238
 Db 4193 AGGATGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 4140
 QY 6239 GTGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 6298
 Db 4139 GCGCGCGACCGACAGCGCGGCTGGGGGCGCTGCGCGCGACAGCTGGCGCGCTGCAAC 4080
 QY 6299 GGTGCGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 6358
 Db 4079 GAGAGCGGCGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 4020
 QY 6359 CTTCGCCAGAGCTGCTTTTGTATCTGATGATGATGATGATGATGATGATGATGATG 6418
 Db 4019 TTCCCGGACCTTTGGCGGATGATGATGATGATGATGATGATGATGATGATGATGATG 3960
 QY 6419 CGTTATGCGGTGTTTGGGGACAGCTGCTTACGCTTACGCTTACGCTTACGCTTACGCT 6478
 Db 3959 GAATACCGGACCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3900
 QY 6479 TTAATGATGCGCGCGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 6538
 Db 3899 GTGATGATGCGCGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3840
 QY 6539 TTGGGCTGCGGCTTTCAGCTTACCAATATGCAAGGATATGATGATGATGATGATGATGAT 6598
 Db 3899 CTGGGCGCTTGGCTTACAGCTGACGAATGCTGCGGAGCTGATGATGATGATGATGATG 3780
 QY 6599 GCGCGTCTATCTGCGCAAAAGCTGCTGATGATGATGATGATGATGATGATGATGATG 6658

Db 3779 GGGCGCTGCTATCTGCTGCTGCGGACCTGCTGCGGAGCGG-----GGGACGCTT 3729
 QY 6659 GCGGATCCGACATGCGCGAGCGCTGCGCGCGCTGCGAGCGCTGTGTCGCGAGGCC 6718
 Db 3728 GAGGATCCGCTGCTGCGAGCGGCTCTATTCGTCATCATCCGCGCTGTGACGCGGCC 3669
 QY 6719 GAGCGGCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 6778
 Db 3668 GAGCGCTATGATGCTGCGCGCGGAGCGGCTTCCGATCTGCGCGCGCTGCGGCTGG 3609
 QY 6779 GCGATGCGACCGCGCGCGGCTTACCGGGAATTGCGGTAAGTGACATGCGGCT 6838
 Db 3608 TCGATGCGCGCGCGCGCTGCTGCTATCTATCGGCAATCGGAGCGGATCCGCGAGGCTG 3549
 QY 6839 GCGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 6898
 Db 3548 CCGAGGCTATGCGCGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3489
 QY 6899 AAAGTCCGCGCTGCTGCTTACTTTCGCGCTTGTCTATCCGAGGCGGCTGCTGCGGCT 6958
 Db 3488 CCGAGGCTTGTGACGCGCGCGATGATGATGATGATGATGATGATGATGATGATGATG 3429
 QY 6959 CTGTCGAGCGTCCGCG 6975
 Db 3428 CTGTGACCCGACCGCG 3412

Search completed: December 8, 2004, 21:45:15
 Job time : 34882 secs